

PCT/JP 99/03929

10.08.99

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

REC'D	22 OCT 1999
WIPO	PCT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

1998年 9月 9日

出願番号
Application Number:

平成10年特許願第254736号

出願人
Applicant(s):

財団法人相模中央化学研究所
株式会社プロテジーン

E3KJ

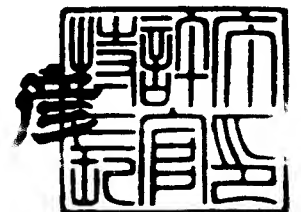
PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

1999年10月 8日

特許庁長官
Commissioner
Patent Office

近藤 隆



出証番号 出証特平11-3059879

【書類名】 特許願

【整理番号】 S018165

【提出日】 平成10年 9月 9日

【あて先】 特許庁長官殿

【発明の名称】 疎水性ドメインを有するヒトタンパク質およびそれをコードするDNA

【請求項の数】 6

【発明者】

 【住所又は居所】 神奈川県相模原市若松 3-4 6-50

 【氏名】 加藤 誠志

【発明者】

 【住所又は居所】 東京都葛飾区高砂 5-1 3-11

 【氏名】 山口 知子

【特許出願人】

 【代表出願人】

 【識別番号】 000173762

 【住所又は居所】 神奈川県相模原市西大沼 4丁目 4番 1号

 【氏名又は名称】 財団法人相模中央化学研究所

 【代表者】 近藤 聖

 【電話番号】 0427(42)4791

【特許出願人】

 【識別番号】 596134998

 【住所又は居所】 東京都目黒区中町 2丁目 20番 3号

 【氏名又は名称】 株式会社プロテジーン

 【代表者】 棚井 文雄

 【電話番号】 03(3792)1019

【手数料の表示】

 【予納台帳番号】 011501

 【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 疎水性ドメインを有するヒトタンパク質およびそれをコードするDNA

【特許請求の範囲】

【請求項1】 配列番号1から配列番号10で表されるアミノ酸配列のいずれかを含む蛋白質。

【請求項2】 請求項1記載の蛋白質のいずれかをコードするDNA。

【請求項3】 配列番号11から配列番号20で表される塩基配列のいずれかを含むcDNA。

【請求項4】 配列番号21から配列番号30で表される塩基配列のいずれかからなる、請求項3記載のcDNA。

【請求項5】 請求項2から請求項4のいずれかに記載のDNAをインビトロ翻訳あるいは真核細胞内で発現しうる発現ベクター。

【請求項6】 請求項2から請求項4のいずれかに記載のDNAを発現し、請求項1記載の蛋白質を生産しうる形質転換真核細胞。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、疎水性ドメインを有するヒト蛋白質、それをコードしているDNA、このDNAの発現ベクター、およびこのDNAを発現させた真核細胞に関する。本発明の蛋白質は、医薬品として、あるいはこの蛋白質に対する抗体を作製するための抗原として用いることができる。本発明のヒトcDNAは、遺伝子診断用プローブや遺伝子治療用遺伝子源として用いることができる。また、このcDNAがコードしている蛋白質を大量生産するための遺伝子源として用いることができる。これらの遺伝子を導入して分泌蛋白質や膜蛋白質を大量発現させた細胞は、対応するレセプターやリガンドの検出、新しい低分子医薬のスクリーニングなどに利用できる。

【0002】

【従来の技術】

細胞は多くの蛋白質を細胞外に分泌している。これらの分泌蛋白質は、細胞の増殖制御、分化誘導、物質輸送、生体防御などにおいて重要な役割を果たしている。分泌蛋白質は細胞内蛋白質と異なり細胞外で作用するので、注射や点滴などによる体内投与が可能であり、医薬としての可能性を秘めている。事実、インターフェロン、インターロイキン、エリスロポイエチン、血栓溶解剤など、多くのヒト分泌蛋白質が現在医薬として使用されている。また、これら以外の分泌蛋白質についても臨床試験が進行中であり、医薬品を目指した用途開発がなされている。ヒト細胞は、まだ多くの未知の分泌蛋白質を生産していると考えられており、これらの分泌蛋白質並びにそれをコードしている遺伝子が入手できれば、これらを用いた新しい医薬品開発が期待できる。

【0003】

一方、膜蛋白質は、シグナルレセプター、イオンチャネル、トランスポーターなどとして、細胞膜を介する物質輸送や情報伝達において重要な役割を担っている。例えば、各種サイトカインに対するレセプター、ナトリウムイオン・カリウムイオン・塩素イオン等に対するイオンチャネル、糖・アミノ酸等に対するトランスポーターなどが知られており、その多くはすでに遺伝子もクローン化されている。これらの膜蛋白質の異常は、これまで原因不明であった多くの病気と関連していることがわかってきた。従って、新しい膜蛋白質が見い出せれば、多くの病気の原因解明につながるものと期待され、膜蛋白質をコードする新たな遺伝子の単離が望まれている。

【0004】

従来、これらの分泌蛋白質や膜蛋白質は、ヒト細胞から精製することが困難なので、遺伝子の方からのアプローチによって単離されたものが多い。一般的な方法は、cDNAライブラリーを真核細胞に導入して、cDNAを発現させたのち、目的とする活性を有する蛋白質を分泌発現あるいは膜表面上に発現している細胞をスクリーニングする、いわゆる発現クローニング法である。しかしこの方法

一般に分泌蛋白質や膜蛋白質は、蛋白質内部に少なくとも一個所疎水性ドメイ

ンを有しており、リボソームで合成された後、このドメインが分泌シグナルとして働いたり、リン脂質膜内に留まり膜にトラップされる。従って、完全長 cDNA の全塩基配列を決定してやり、その cDNA がコードしている蛋白質のアミノ酸配列の中に疎水性の高い領域が存在すれば、その cDNA は分泌蛋白質や膜蛋白質をコードしていると考えられる。

【0006】

【発明が解決しようとする課題】

本発明の目的は、疎水性ドメインを有する新規のヒト蛋白質、この蛋白質をコードする DNA、この DNA の発現ベクター、およびこの DNA を発現しうる形質転換真核細胞を提供することである。

【0007】

【課題を解決するための手段】

本発明者らは鋭意研究の結果、ヒト完全長 cDNA バンクの中から疎水性ドメインを有する蛋白質をコードする cDNA をクローン化し、本発明を完成した。すなわち、本発明は疎水性ドメインを有するヒト蛋白質である、配列番号 1 から配列番号 10 で表されるアミノ酸配列のいずれかを含む蛋白質を提供する。また本発明は上記蛋白質をコードする DNA、例えば配列番号 11 から配列番号 30 で表される塩基配列のいずれかを含む cDNA、並びにこの DNA をインビトロ翻訳あるいは真核細胞内で発現しうる発現ベクター、及びこの DNA を発現し上記蛋白質を生産しうる形質転換真核細胞を提供する。

【0008】

【発明の実施の形態】

本発明の蛋白質は、ヒトの臓器、細胞株などから単離する方法、本発明のアミノ酸配列に基づき化学合成によってペプチドを調製する方法、あるいは本発明の疎水性ドメインをコードする DNA を用いて組換え DNA 技術で生産する方法などにより取得することができるが、組換え DNA 技術で取得する方法が好ましく用いられる。例えば、本発明の cDNA を有するベクターからインビトロ転写によって RNA を調製し、これを鋳型としてインビトロ翻訳を行なうことによりインビトロで蛋白質を発現できる。また翻訳領域を公知の方法により適当な発現ベ

クターに組換えてやれば、大腸菌、枯草菌等の原核細胞や、酵母、昆虫細胞、哺乳動物細胞等の真核細胞で、コードしている蛋白質を大量に発現させることができる。

【0009】

本発明の蛋白質を、インビトロ翻訳でDNAを発現させて生産させる場合には、このcDNAの翻訳領域を、RNAポリメラーゼプロモーターを有するベクターに組換え、プロモーターに対応するRNAポリメラーゼを含む、ウサギ網状赤血球溶解物や小麦胚芽抽出物などのインビトロ翻訳系に添加してやれば、本発明の蛋白質をインビトロで生産することができる。RNAポリメラーゼプロモーターとしては、T7、T3、SP6などが例示できる。これらのRNAポリメラーゼプロモーターを含むベクターとしては、pKA1、pCDM8、pT3/T7 18、pT7/3 19、pBluescript IIなどが例示できる。また、反応系にイヌ臍臓ミクロソームなどを添加してやれば、本発明の蛋白質を分泌型あるいはミクロソーム膜に組み込まれた形で発現することができる。

【0010】

本発明の蛋白質を、大腸菌などの微生物でDNAを発現させて生産させる場合には、微生物中で複製可能なオリジン、プロモーター、リボソーム結合部位、cDNAクローニング部位、ターミネーター等を有する発現ベクターに、本発明のcDNAの翻訳領域を組換えた発現ベクターを作成し、この発現ベクターで宿主細胞を形質転換したのち、得られた形質転換体を培養してやれば、このcDNAがコードしている蛋白質を微生物内で大量生産することができる。この際、任意の翻訳領域の前後に開始コドンと停止コドンを付加して発現させてやれば、任意の領域を含む蛋白質断片を得ることができる。あるいは、他の蛋白質との融合蛋白質として発現させることもできる。この融合蛋白質を適当なプロテアーゼで切断することによってこのcDNAがコードする蛋白質部分のみを取得することもできる。大腸菌用発現ベクターとしては、pUC系、pBluescript

本発明の蛋白質を、真核細胞でDNAを発現させて生産させる場合には、この

cDNAの翻訳領域を、プロモーター、スプライシング領域、ポリ(A)付加部位等を有する真核細胞用発現ベクターに組換え、真核細胞内に導入してやれば、本発明の蛋白質を分泌生産あるいは膜蛋白質として細胞膜表面上で生産することができる。発現ベクターとしては、pKA1、pED6dpc2、pCDM8、pSVK3、pMSG、pSVL、pBK-CMV、pBK-RSV、EBVベクター、pRS、pYES2などが例示できる。真核細胞としては、サル腎臓細胞COS7、チャイニーズハムスター卵巣細胞CHOなどの哺乳動物培養細胞、出芽酵母、分裂酵母、カイコ細胞、アフリカツメガエル卵細胞などが一般に用いられるが、本蛋白質を発現できるものであれば、いかなる真核細胞でもよい。発現ベクターを真核細胞に導入するには、電気穿孔法、リン酸カルシウム法、リボソーム法、DEAEデキストラン法など公知の方法を用いることができる。

【0012】

本発明の蛋白質を原核細胞や真核細胞で発現させたのち、培養物から目的蛋白質を単離精製するためには、公知の分離操作を組み合わせで行うことができる。例えば、尿素などの変性剤や界面活性剤による処理、超音波処理、酵素消化、塩析や溶媒沈殿法、透析、遠心分離、限外濾過、ゲル濾過、SDS-PAGE、等電点電気泳動、イオン交換クロマトグラフィー、疎水性クロマトグラフィー、アフィニティークロマトグラフィー、逆相クロマトグラフィーなどがあげられる。

【0013】

本発明の蛋白質には、配列番号1から配列番号10で表されるアミノ酸配列のいかなる部分アミノ酸配列を含むペプチド断片(5アミノ酸残基以上)も含まれる。これらのペプチド断片は抗体を作製するための抗原として用いることができる。また、本発明の蛋白質の中でシグナル配列を有するものは、シグナル配列が除去された後、成熟蛋白質の形で分泌される。したがって、これらの成熟蛋白質は本発明の蛋白質の範疇にはいる。成熟蛋白質のN末端アミノ酸配列は、シグナル配列切断部位決定法[特開平8-187100]を用いて容易に求めることができる。また、いくつかの膜蛋白質は、細胞表面でプロセシングを受けて分泌型となる。このような分泌型となった蛋白質あるいはペプチドも本発明の蛋白質の範疇にはいる。アミノ酸配列の中に糖鎖結合部位が存在すると、適当な真核細胞

で発現させれば糖鎖が付加した蛋白質が得られる。したがって、このような糖鎖が付加した蛋白質あるいはペプチドも本発明の蛋白質の範疇にはいる。

【0014】

本発明のDNAには、上記蛋白質をコードするすべてのDNAが含まれる。このDNAは、化学合成による方法、cDNAクローニングによる方法などを用いて取得することができる。

【0015】

本発明のcDNAは、例えばヒト細胞由来cDNAライブラリーからクローン化することができる。cDNAはヒト細胞から抽出したポリ(A)⁺RNAを鋳型として合成する。ヒト細胞としては、人体から手術などによって摘出されたものでも培養細胞でも良い。cDNAは、岡山-Berg法[Okayama, H. and Berg, P., Mol. Cell. Biol. 2:161-170 (1982)]、Gubler-Hoffman法[Gubler, U. and Hoffman, J., Gene 25:263-269 (1983)]などいかなる方法を用いて合成してもよいが、完全長クローンを効率的に得るためには、実施例にあげたようなキャッピング法[Kato, S. et al., Gene 150:243-250 (1994)]を用いることが望ましい。また市販のヒトcDNAライブラリーを用いることもできる。cDNAライブラリーから本発明のcDNAをクローン化するには、本発明のcDNAの任意の部分の塩基配列に基づいてオリゴヌクレオチドを合成し、これをプローブとして用いて、公知の方法によりコロニーあるいはブラークハイブリダイゼーションによるスクリーニングを行えばよい。また、目的とするcDNA断片の両末端にハイブリダイズするオリゴヌクレオチドを合成し、これをプライマーとして用いて、ヒト細胞から単離したmRNAからRT-PCR法により、本発明のcDNA断片を調製することもできる。

【0016】

本発明のcDNA断片は、その塩基配列が、本発明のcDNA断片の塩基配列と一致するものである。それぞれのクローン番号(HP番号)、cDNAクロー

ンが得られた細胞、cDNAの全塩基数、コードしている蛋白質のアミノ酸残基数をそれぞれ表1にまとめて示した。

【0017】

【表1】

表1

配列番号	HP番号	細胞	塩基数	アミノ酸 残基数
1、11、21	HP01470	胃癌	1619	358
2、12、22	HP02419	胃癌	2054	226
3、13、23	HP02631	Sa o s - 2	1380	195
4、14、24	HP02695	胃癌	1292	339
5、15、25	HP10031	Sa o s - 2	2168	487
6、16、26	HP10530	Sa o s - 2	1357	393
7、17、27	HP10541	胃癌	711	196
8、18、28	HP10550	胃癌	651	107
9、19、29	HP10590	HT-1080	1310	350
10、20、30	HP10591	HT-1080	1400	107

【0018】

なお、配列番号11から配列番号30のいずれかに記載のcDNAの塩基配列に基づいて合成したオリゴヌクレオチドプローブを用いて、本発明で用いたヒト細胞株やヒト組織から作製したcDNAライブラリーをスクリーニングすることにより、本発明のcDNAと同一のクローンを容易に得ることができる。

【0019】

一般にヒト遺伝子は個体差による多型が頻繁に認められる。従って配列番号11から配列番号30において、1又は複数個のヌクレオチドの付加、欠失および／又は他のヌクレオチドによる置換がなされているcDNAも本発明の範疇には

いる。

【0020】

同様に、これらの変更によって生じる、1又は複数個のアミノ酸の付加、欠失および／又は他のアミノ酸による置換がなされている蛋白質も、配列番号1から配列番号10で表されるアミノ酸配列を有するそれぞれの蛋白質の活性を有する限り、本発明の範疇に入る。

【0021】

本発明のcDNAには、配列番号11から配列番号20で表される塩基配列あるいは配列番号21から配列番号30で表される塩基配列のいかなる部分塩基配列を含むcDNA断片(10bp以上)も含まれる。また、センス鎖およびアンチセンス鎖からなるDNA断片もこの範疇にはいる。これらのDNA断片は遺伝子診断用のプローブとして用いることができる。

【0022】

【実施例】

次に実施例により発明を具体的に説明するが、本発明はこれらの例に限定されるものではない。DNAの組換えに関する基本的な操作および酵素反応は、文献["Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Laboratory, 1989]に従った。制限酵素および各種修飾酵素は特に記載の無い場合宝酒造社製のものをを用いた。各酵素反応の緩衝液組成、並びに反応条件は付属の説明書に従った。cDNA合成は文献[Kato, S. et al., Gene 150:243-250 (1994)]に従った。

【0023】

(1) 疎水性ドメインを有する蛋白質をコードしているcDNAの選別

cDNAライブラリーとして、フィブロサルコーマ細胞株HT-1080 cDNAライブラリー(WO98/11217)、骨肉腫細胞株Saos-2 cDN

ら完全長cDNAクローンを選択し、その全塩基配列決定を行い、完全長cDN

Aクローンからなるホモ・プロテインcDNAバンクを構築した。ホモ・プロテインcDNAバンクに登録された完全長cDNAクローンがコードしている蛋白質について、Kyte-Doolittleの方法[Kyte, J & Doolittle, R. F., J. Mol. Biol. 157:105-132 (1982)]により、疎水性／親水性プロフィールを求め、疎水性ドメインの有無を調べた。コードしている蛋白質のアミノ酸配列中に分泌シグナルや膜貫通ドメインと思われる疎水的な領域があるクローンを候補クローンとして選別した。

【0024】

(2) インビトロ翻訳による蛋白質合成

本発明のcDNAを有するプラスミドベクターを用いて、T_NTウサギ網状赤血球溶解物キット（プロメガ社製）によるインビトロ転写／翻訳を行なった。この際 [³⁵S] メチオニンを添加し、発現産物をラジオアイソトープでラベルした。いずれの反応もキットに付属のプロトコールに従って行なった。プラスミド2 μ gを、T_NTウサギ網状赤血球溶解物12.5 μ l、緩衝液（キットに付属）0.5 μ l、アミノ酸混合液（メチオニンを含まない）2 μ l、 [³⁵S] メチオニン（アマーシャム社）2 μ l（0.37MBq/ μ l）、T7RNAポリメラーゼ0.5 μ l、RNasin20Uを含む総量25 μ lの反応液中で30℃で90分間反応させた。また、膜系存在下の実験は、この反応系に、イヌ脾臓ミクロソーム画分（プロメガ）2.5 μ lを添加して行なった。反応液3 μ lにSDSサンプリングバッファー（125mMトリス塩酸緩衝液、pH6.8、120mM2-メルカプトエタノール、2%SDS溶液、0.025%ブロモフェノールブルー、20%グリセロール）2 μ lを加え、95℃3分間加熱処理した後、SDS-ポリアクリルアミドゲル電気泳動にかけた。オートラジオグラフィーを行ない、翻訳産物の分子量を求めた。

【0025】

(3) COS7による発現

本発明の蛋白質の発現ベクターを有する大腸菌を100 μ g/mlアンピシリン含有2xYT培地2ml中で37℃2時間培養した後、ヘルパーファージM13KO7（50 μ l）を添加し、37℃で一晩培養した。遠心によって分離した

上澄からポリエチレングリコール沈殿によって一本鎖ファージ粒子を得た。これを100 μ lの1 mM トリス-0.1 mM EDTA、pH 8 (TE) に懸濁した。

【0026】

サル腎臓由来培養細胞COS 7は、10%ウシ胎児血清を含むダルベッコ改変イーグル (DMEM) 培地中、5%CO₂存在下、37℃で培養した。1 x 10⁵個のCOS 7細胞を6穴プレート (ヌンク社、穴の直径3 cm) に植え、5%CO₂存在下、37℃で22時間培養した。培地除去後、リン酸緩衝液で細胞表面を洗浄し、さらに50 mM トリス塩酸 (pH 7.5) を含むDMEM (TDMEM) で再度洗浄した。この細胞に一本鎖ファージ懸濁液1 μ l、DMEM 培地0.6 ml、TRANSFECTAMTM (IBF社) 3 μ lを懸濁したものを添加し、5%CO₂存在下、37℃で3時間培養した。サンプル液を除去後、TDMEMで細胞表面を洗浄し、10%ウシ胎児血清含有DMEMを1穴あたり2 ml 加え、5%CO₂存在下、37℃にて2日間培養した。培地を [³⁵S] システインあるいは [³⁵S] メチオニンを含む培地に交換した後、1時間培養した。遠心分離によって、培地と細胞を分けたあと、培地画分と細胞膜画分の蛋白質をSDS-PAGEにかけた。

【0027】

(4) クローン例

<HP01470> (配列番号1、11、21)

ヒト胃癌cDNAライブラリーから得られたクローンHP01470のcDNAインサートの全塩基配列を決定したところ、157 bpの5' 非翻訳領域、1077 bpのORF、385 bpの3' 非翻訳領域からなる構造を有していた。ORFは358アミノ酸残基からなる蛋白質をコードしており、1箇所の推定膜貫通ドメインが存在した。図1にKyte-Doolittleの方法で求めた本蛋白質の疎水性/親水性プロフィールを示す。インビトロ翻訳の結果、ORF

際、
る40 kDaと何らかの修飾を受けたと思われる43.5 kDaの産物が生成し

た。分泌シグナル配列切断部位予測法である（-3、-1）規則を適用すると、成熟蛋白質は23番目のグリシンから始まると予想される。

【0028】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、線虫仮想蛋白質39.9kDa（SWISS-PROTアクセション番号Q10005）と類似性を有していた。表2に、本発明のヒト蛋白質（HP）と線虫仮想蛋白質39.9kDa（CE）のアミノ酸配列の比較を示す。-はギャップを、*は本発明の蛋白質と同一アミノ酸残基を、. は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。全領域にわたって、58.9%の相同性を有していた。

【0029】

【表2】

表2

HP	MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRASIKDIKKAYRKLALQLHPDRNPDD
	*. . * *****. . . * . . ***** *****.*
CE	MRILNVSLVLASSLVAFVECGRDFYKILGVAKNANANQIKKAYRKLAKELHPDRNQDD
HP	PQAQEFQDLGAAYEVLSDSEKRRKYDITYGEEGL--KDGHQSSHGDIFSHFFGDFGFMFG
	*.*****.*****.* ** .****. . . * . . * ** ***** *
CE	EMANEKFQDLSSAYEVLSDKEKRAMYDRHGEEGVAKMGGGGGGHDPFSSFFGDF-FG-G
HP	GTPRQQDRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTT
 * *****.*.***. *.*. **.*****.****.
CE	GGHGGEETPKGADVTIDLFVTLEEVYNGHFVEIKRKKAVYKQTSQTRQCNCRHEMRTE
HP	QLGPGRFQMTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMETPFIGEGEPHVDGEPGD
	*. . ***** * *****.*.*. * . . * *****.*.***
CE	QMGQGRFQMFQVVCDECPNVKLVQENKVLEVEVEVGADNGHQIFHGEGEPHIEGDPGD
HP	LRFKIRIKVVKHPIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVVHISRDKITRPGAK
	*. * ** * * * * * *
CE	LKFKIRIQKHPRFERKGDDLYTNVTISLQDALNGFEMEIQHLDGHIVKVQRDKVTWPGAR
HP	LWKKGEGLPNFDNNNIKGSIIITFDVDFPKEQLTEEAREGIKQLLKQGSVQ-KVYNGLQG

*.***.***.*... ** ** *.***.***.*... *...*.***.*.***

CE LRKKDEGMPSLEDNNKKGMLVVTFDVEFPKTELSDEQKAQIIEILQQNTVKPKAYNGL

【 0 0 3 0 】

また、本 cDNA の塩基配列を用いて G e n B a n k を検索したところ、E S T の中に、90%以上の相同性を有するもの（例えば、アクセシオン番号 A A 2 8 2 8 3 8）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【 0 0 3 1 】

<HP02419> (配列番号 2、12、22)

ヒト胃癌 cDNA ライブラリーから得られたクローン HP02419 の cDNA インサートの全塩基配列を決定したところ、253bp の 5' 非翻訳領域、681bp の ORF、1120bp の 3' 非翻訳領域からなる構造を有していた。ORF は 226 アミノ酸残基からなる蛋白質をコードしており、4 箇所の推定膜貫通ドメインが存在した。図 2 に Kyte-Doolittle の方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インビトロ翻訳の結果、高分子量の翻訳産物が生成した。

【 0 0 3 2 】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ヒト仮想蛋白質K I A A 0 1 0 8 (S W I S S - P R O Tアクセス番号Q 1 5 0 1 2)と類似性を有していた。表3に、本発明のヒト蛋白質(HP)とヒト仮想蛋白質K I A A 0 1 0 8 (K I)のアミノ酸配列の比較を示す。－はギャップを、*は本発明の蛋白質と同一アミノ酸残基を、. は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。全領域にわたって、43.9%の相同性を有していた。

【 0 0 3 3 】

【表 3】

HP MKMVAPWTRFYSNSCCLCCHVRTGTILLGVWYLIINAVVLLILLSALADPD---QY

****. ** *****. **. **. . . . * *

KI MVSMFKNRSDRFYSTRCCGCHVRTGTIILGTWYMVVNLLMAILLTVETHPN SMPAV
HP NFSSSELGGDFEF-MDDANMCIAIAISLLMILICAMATYGAYKQRAAWIIPFFCYQIFDF

*. . . *. * *. . . *. * * *

KI NIQYEVIGNYYSSERMADNACVLFAVSVLMFISSMLVYGAISYQVGWLIPFFCYRLDF
HP ALNMLVAITVLIYPNSIQEYIRQLPPNFPYRDDVMSVNPTCLVLIILLFISILTFKGYL

. *. ****. *. * . . . *. *. ** *. ****. *. *. *. *

KI VLSCLVAISSTYLPRIKEYLDQL-PDFPYKDDLALDSSCLLFIVLVFFALFIIFKAYL
HP ISCVWNCYRYINGRNSSDVLVYVT-SNDTTVLLPPYDDATVNGAAKEPPPPYVSA

*. ****. *. *. . . . * * * *

KI INCVWNCYKYINNRNVPEIAVYPAFEAPPQYVLPTY-EMAVKMPEKEPPPPYLP

【0034】

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセション番号AA173214）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【0035】

<HP02631>（配列番号3、13、23）

ヒト骨肉腫細胞株Saos-2cDNAライブラリーから得られたクローンHP02631のcDNAインサートの全塩基配列を決定したところ、42bpの5'非翻訳領域、588bpのORF、750bpの3'非翻訳領域からなる構造を有していた。49番目のアミノ酸残基は停止コドンになっているが、翻訳産物の分子量や線虫ホモログとの配列比較データを参照すると、この部分はセレノシステインである可能性が高い。ORFは195アミノ酸残基からなる蛋白質をコードしており、N末端に推定分泌シグナルが、中間部に1箇所の推定膜貫通ドメインが存在した。図3にKyte-Doolittleの方法で求めた本蛋白質の疎水性/親水性プロファイルを示す。インビトロ翻訳の結果、58kDaの翻訳産物が生成した。この際、ミクロソームを添加すると、分泌シグナルが切断

された考えられる 56 kDa の産物が生成した。いずれも、ORF から予想される分子量 22 kDa より大きいので、他の蛋白質と相互作用している可能性が高い。

【0036】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、線虫仮想蛋白質 C35C5.3 (EMBL アクセション番号 Z78417) と類似性を有していた。表 4 に、本発明のヒト蛋白質 (HP) と線虫仮想蛋白質 C35C5.3 (CE) のアミノ酸配列の比較を示す。本蛋白質のアミノ酸配列の中で、49 番目の U はセレノシステインを表す。- はギャップを、* は本発明の蛋白質と同一アミノ酸残基を、. は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。N 末端領域を除く全領域にわたって、37.9% の相同性を有していた。本蛋白質の 49 番目の停止コドン (セレノシステイン) の部分は、線虫蛋白質においてはシステインとなっていた。

【0037】

【表 4】

表 4

HP	MRLLLL
CE MRIHDELQKQDMSRFGVFIIGVLFFMSVCDVLRTEESHSHDENHVHEKDDFEAEFGDETDS	
HP LLVAASAMVRSEASANLGGVPSKRLKMQYATGPLLKFKICVSUGYRRVFEEYMRVISQRY	
	* * . . . * * * * *
CE QSFSQGTEEDHIEVREQSSFKPTAVHHAKDLPTLRIFYCVSCGYKQAFDQFTTFAKEY	
HP PDIRIEGENYLPQPIYRHIAFLSVFKLVLIIGLIIVGKDPFAFFGMQAPSIWQWQENKV	
	* . . . * * * * * * *
CE PNMPIEGANFAPVLWKAYVAQALSFKMAVLVLVGGINPFRFGLGYPQILQHAHGKMK	
CE SSCMLVFMLGNLVEQSLISTGAFFVYLGNEQIWSKIESGRVPSPQEFMQLIDAQLAVLGK	

HP MDSIPHRS

CE APVNTESFGFEQQTV

【0038】

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセシオン番号AA156969）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【0039】

<HP02695>（配列番号4、14、24）

ヒト胃癌cDNAライブラリーから得られたクローンHP02695のcDNAインサートの全塩基配列を決定したところ、112bpの5'非翻訳領域、1020bpのORF、160bpの3'非翻訳領域からなる構造を有していた。ORFは339アミノ酸残基からなる蛋白質をコードしており、3箇所の推定膜貫通ドメインが存在した。図4にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロファイルを示す。インビトロ翻訳の結果、ORFから予想される分子量38,274とほぼ同じ38kDaの翻訳産物が生成した。

【0040】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ラット高血圧誘導蛋白質S-2断片（PIRアクセシオン番号539959）と類似性を有していた。表5に、本発明のヒト蛋白質（HP）とラット高血圧誘導蛋白質S-2断片（RN）のアミノ酸配列の比較を示す。-はギャップを、*は本発明の蛋白質と同一アミノ酸残基を、.は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。全領域にわたって、74.3%の相同性を有していた。

【0041】

【表5】

表5

HP MNWELLLWLLVLCALLLLLVQLLRFLRADGDLTLLWAEWQGRRPEWELTDMVVVWTGASS

HP GIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVPLDLTDTGSHEA

****.*****.***.***.

RN VKRRSLENGNLKEKDILVPLDLADTSSHDI

HP ATKAVLQEFGRIDILVNNGGMSQRSCLMDTSLDVYRKLIENYLGTVSLTKCVLPHMIER

.**... ** .*.***.***.***** ****.***

RN ATKTVLQEFGRIDILVNNGGVAHASLVENTNMDIFKVLIEVNYLGTVSLTKCFLPHMMER

HP KQGKIIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIIVSNICPGPVQSN

.*****...*

RN NQGKIIVMKKS

【0042】

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセシオン番号T84331）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【0043】

<HP10031>（配列番号5、15、25）

ヒト骨肉腫細胞株Saos-2cDNAライブラリーから得られたクローンHP10031のcDNAインサートの全塩基配列を決定したところ、55bpの5'非翻訳領域、1464bpのORF、649bpの3'非翻訳領域からなる構造を有していた。ORFは487アミノ酸残基からなる蛋白質をコードしており、11箇所の推定膜貫通ドメインが存在した。図5にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロファイルを示す。インビ

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、

線虫仮想蛋白質 CELK07H8 (GenBank アクセション番号 AF047659) と類似性を有していた。表 6 に、本発明のヒト蛋白質 (HP) と線虫仮想蛋白質 CELK07H8 (CE) のアミノ酸配列の比較を示す。- はギャップを、* は本発明の蛋白質と同一アミノ酸残基を、. は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。N 末端側を除く全領域にわたって、44.2% の相同性を有していた。

【0045】

【表 6】

表 6

HP	MDGTETRQRRLDSCGKPGELGLPHPLSTGGLPVAS
CE	MKGGGGIGDGKKDYQSAVHEGLTTFDQLGIALEDVGKSMDAETATPGGSLFSRVIFRFRN
HP	EDGALRAPESQSVTPKPLETEPSRETAWSIGLQVTVPFMAFAGLGLSWAGMLLDYFQHWPV
 *... . ** ** ****. **... **.
CE	ENSSLKSRTYDHSNDLVNMSVIPAESSVLFQVLPFAVAGLGMVFAGLVLSIVVTWPL
HP	FVEVKDLLTLVPPLVGLKGNLEMTLASRLSTAANTGQIDDPQEQHRVISSNLALIQVQAT
	* * . . *.****.*.***** ** *..*..... *. *****.*****
CE	FEEIPEILILVPALLGLKGNLEMTLASRLSTLANLGHMDSSKQRKDVVIANLALVQVQAT
HP	VVGLLAAVAALLLVGSREEVDVAKVELLCASSVLTAFLLAALGVLVCIVIGARKLGV
 * *. * *. *.****. ** *...*.***** ..
CE	VVAFLASAFAAALAFIPSGDFDWAHGALMCASSLATACSASLVLSLLMVVIVTSRKYN I
HP	NPDNIATPIAASLGLDITLSILALVSSFFYR-HKDSRYLTPLVCLSFAALTPVWVLIQKQ
	****.*****.***.***. . * * . *.....* . * . * * * . **..
CE	NPDNVATPIAASLGLDTTLTVLAFFGSVFLKAHNTESWLNVIIVIVLFLLLLPFWIKIANE
HP	SPPIVKILKFGWFPIILAMVISSFGGLILSKTVSKQYKGM AIFTPVICGVGGNLVAIQT
	. . . * ** *.***.*** **..***.*** ..*.....**. *****.*.*.
CE	NEGTQETLYNGWTPVIMSMLISSAGGFILETAV--RRYHSLSTYGPVLNGVGGNLAAVQA
HP	SRISTYLHMWSAPGVLP LQ--MKKFWPNPCSTFCTSEINMSARVLLLLVVPGLHIF-FY

[illegible]

【 0 0 4 6 】

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセシオン番号AA334000）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【 0 0 4 7 】

<HP10530> (配列番号 6、16、26)

ヒト骨肉腫細胞株 S a o s - 2 c D N A ライブラリーから得られたクローン H P 1 0 5 3 0 の c D N A インサートの全塩基配列を決定したところ、80 b p の 5' 非翻訳領域、1182 b p の O R F、95 b p の 3' 非翻訳領域からなる構造を有していた。O R F は 393 アミノ酸残基からなる蛋白質をコードしており、N 末端に推定分泌シグナルが存在した。図 6 に K y t e - D o o l i t t l e の方法で求めた本蛋白質の疎水性／親水性プロファイルを示す。インビトロ翻訳の結果、O R F から予想される分子量 44,912 よりやや大きい 46 k D a の翻訳産物が生成した。この際、ミクロソームを添加すると、分泌シグナルが切断された考えられる 45.5 k D a の産物が生成した。分泌シグナル配列切断部位予測法である (-3, -1) 規則を適用すると、成熟蛋白質は 23 番目のリジン

1000

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、

シロイナズナ仮想蛋白質 IG002N01 (GenBank アクセション番号 AF007269) と類似性を有していた。表 7 に、本発明のヒト蛋白質 (HP) とシロイナズナ仮想蛋白質 IG002N01 (AT) のアミノ酸配列の比較を示す。- はギャップを、* は本発明の蛋白質と同一アミノ酸残基を、. は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。N 末端側 355 アミノ酸残基にわたって、27.0% の相同性を有していた。

【0049】

【表 7】

表 7

HP	MRTLFLNLLWL
AT	MELTSFQKSPSSNDVVSFSVSLVRNSMARRRRSSAAESLKRRNDGYESLCQVVQQDSDRR
HP	ALACSPVHTTSLKSDAKKAASKTLLEKSQFSDKPVQDRGLVVTDLKAESVVLEHRSYCSA
*.* **.. **.. ..
AT	LITIFVIFVIVIPAVSIAVYKVKFADRVIQTSSIRQKGIVKTDINFQEILTEHSK--AS
HP	KARDRHFAAGDVLGYVTPWNSHGYDVTKVFGSKFTQISPVWLQ-LKRRRGREMFVETGLHDV
**.. *.**.*. *..* .. *.... . . *.** *..**.... . **....
AT	ENSTRHYDYPVLAYITP--CQGSG--VLEGR-HNADKGWIQELRSRGNALSASKGLPKL
HP	DQGWMAVRKHAKGLHIVPRLLFEDWTYDDFRNVLDSEDEIEELSKTVVQVAKNQHFDFG
 * . . . * . ** .. *.** *.**
AT	---YNSCIFHALKRMNFFTLELVNFNTYLVIMFALNS-REMEYNGIVLESWSRWAAYGVL
HP	VVEVWNQLLSQKRVGLIHMLTHLAEALHQAARLLALLVIPPAITPGTDQLGMFTHKEFEQL
 * . * * . * *
AT	HDPDLRKMALKFVKQLGDALHSTSSPRNNQQHMQFMVVGPPRSEKLQMYDFGPEDLQFL
HP	APVLDGFSMLTYDYSTAHPGPNAPLSWVRACVQ-VLDPKSK----WRSKILLGLNFYGM
	.*****.*.....*****.*. .. *.***** .*.***.*****
AT	KDSVDGFSLMTYDFSNPQNPQGNAPVWKIDLTLLKLLGSSNNIDSNIARKVLLGINFYGN
HP	DYATSKDAREPVGARYIQTLKDHPRMVWDSQASEHFFEYKKSRSGRHVVFYPTLKSQ

...* *..*...* . **...**.* *... ..*.....*.
 AT DFVISGGGGGAITGRDYLALLQKHKPTFRWDKESGEHLFMYRDDKNIKHAVFYPTLMSIL
 HP VRLELARELGVGVSIIWELGQGLDYFYDLL

.*** ** *.**.....**..*
 AT LRLENARLWGIGISIIWEIGQDKGHFGKYAEASLEASSIFSGHTFDMQFRTNPRQLSRNGS

【0050】

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセシオン番号AA302913）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【0051】

<HP10541>（配列番号7、17、27）

ヒト胃癌cDNAライブラリーから得られたクローンHP10541のcDNAインサートの全塩基配列を決定したところ、7bpの5'非翻訳領域、591bpのORF、113bpの3'非翻訳領域からなる構造を有していた。ORFは196アミノ酸残基からなる蛋白質をコードしており、N末端に推定分泌シグナルが存在した。図7にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インビトロ翻訳の結果、ORFから予想される分子量21,553よりやや大きい23kDaの翻訳産物が生成した。この際、ミクロソームを添加すると、分泌シグナルが切断された考えられる20kDaと糖鎖が付加されたと考えられる23kDaの産物が生成した。分泌シグナル配列切断部位予測法である（-3、-1）規則を適用すると、成熟蛋白質は41番目のグリシンから始まると予想される。なお、この蛋白質のアミノ酸配列の中には、N-グリコシレーションが起こる可能性がある部位が1箇所（185番目Asn-Leu-Thr）存在する。

本蛋白質の... 酸配列...
 ヒトチモーゲン膜蛋白質（GenBankアクセシオン番号AF056492）

と類似性を有していた。表 8 に、本発明のヒト蛋白質（HP）とヒトチモーゲン膜蛋白質（ZM）のアミノ酸配列の比較を示す。－はギャップを、＊は本発明の蛋白質と同一アミノ酸残基を、．は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。C末端側 133 アミノ酸残基にわたって、37.6%の相同性を有していた。

【0053】

【表 8】

表 8

HP MWRVPGTTRRPVTGESPGMHRPEAMLLLLTLALLGGPTWAGKMYGPGGGKYFS-TTEDYD

.** ** *

ZM MLTVALLALLCASASGNAIQARSSSYSGEYGS GGGKRFSHSGNQLD

HP HEITGLRVSVGLLLVKSQVVKLGDSWDVKLGALGGNTQEVTLPGEYITKVFVAFQAFLR

.**.*. . .**. *. *. *. *. *. *. *.**** .. .*.

ZM GPITALRVRVNTYYIVGLQVRYGKVWSDYVGGRNGDLEEIFLHPGESVIQVSGKYKWYLK

HP GMVMTSKDRYFYFGKLDGQISSAYPSQEGQVLVGIYGQYQLLGIKSIGFEWN-YPLEEP

.*. *.****. *** .*. *. *. . .** ***. * *.****.*. **

ZM KLVFVTDKGRYLSFGKDSGTSFNAVPLHPNTVLRFI SGRSGSL-IDAIGLHWDVYPTSCS

HP TTEPPVNLTYSANSPVGR

ZM RC

【0054】

また、本 cDNA の塩基配列を用いて GenBank を検索したところ、EST の中に、90%以上の相同性を有するもの（例えば、アクセシオン番号 AA340605）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【0055】

<HP10550>（配列番号 8、18、28）

ヒト胃癌cDNAライブラリーから得られたクローンHP10550のcDNAインサートの全塩基配列を決定したところ、241bpの5'非翻訳領域、324bpのORF、86bpの3'非翻訳領域からなる構造を有していた。ORFは107アミノ酸残基からなる蛋白質をコードしており、1箇所の推定膜貫通ドメインが存在した。図8にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロファイルを示す。インビトロ翻訳の結果、高分子量の翻訳産物が生成した。

【0056】

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセシオン番号AA348310）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【0057】

<HP10590>（配列番号9、19、29）

ヒトフィブ्रोサルコーマ細胞株HT-1080cDNAライブラリーから得られたクローンHP10590のcDNAインサートの全塩基配列を決定したところ、77bpの5'非翻訳領域、1053bpのORF、180bpの3'非翻訳領域からなる構造を有していた。ORFは350アミノ酸残基からなる蛋白質をコードしており、1箇所の推定膜貫通ドメインが存在した。図9にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロファイルを示す。インビトロ翻訳の結果、ORFから予想される分子量39,285とほぼ同じ40kDaの翻訳産物が生成した。この際、ミクロソームを添加すると、糖鎖が付加されたと考えられる43kDaの産物が生成した。なお、この蛋白質のアミノ酸配列の中には、N-グリコシレーションが起こる可能性がある部位が2箇所（144番目Asn-Asn-Ser、328番目Asn-Leu-Thr）存在する。

また、本cDNAの塩基配列を用いてGenBankを検索したところ、

ESTの中に、90%以上の相同性を有するもの（例えば、アクセシオン番号AA4

61346) が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【0059】

<HP10591> (配列番号 10、20、30)

ヒトフィブ्रोサルコーマ細胞株 HT-1080 cDNA ライブラリーから得られたクローン HP10591 の cDNA インサートの全塩基配列を決定したところ、232bp の 5' 非翻訳領域、324bp の ORF、844bp の 3' 非翻訳領域からなる構造を有していた。ORF は 107 アミノ酸残基からなる蛋白質をコードしており、1箇所の推定膜貫通ドメインが存在した。図 10 に Kyte-Doolittle の方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インビトロ翻訳の結果、ORF から予想される分子量 11,328 とほぼ同じ 12kDa の翻訳産物が生成した。

【0060】

また、本 cDNA の塩基配列を用いて GenBank を検索したところ、EST の中に、90% 以上の相同性を有するもの（例えば、アクセッション番号 H09424）が登録されていたが、部分配列なので本発明の蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

【0061】

【発明の効果】

本発明は疎水性ドメインを有するヒト蛋白質、それをコードしている DNA、この DNA の発現ベクター、およびこの DNA を発現させた真核細胞を提供する。本発明の蛋白質は、いずれも分泌されるかあるいは細胞膜に存在するので、細胞の増殖や分化を制御している蛋白質と考えられる。したがって、本発明の蛋白質は、細胞の増殖や分化の制御に関わる制癌剤などの医薬品として、あるいはこの蛋白質に対する抗体を作製するための抗原として用いることができる。本発明の DNA は、遺伝子診断用プローブや遺伝子治療用遺伝子源として用いることができる。また、この DNA を用いることにより、この蛋白質を大量に発現することができる。これら遺伝子を導入してこの蛋白質を発現させた細胞は、対応するレセプターやリガンドの検出、新しい低分子医薬のスクリーニングなどに利用で

きる。

【0062】

【配列表】

<110> Sagami Chemical Reserach Center,

Protegene Inc.

<120> Human proteins having hydrophobic domains and DNAs encoding these proteins

<130>S018165

<160> 30

【0063】

<210> 1

<211> 358

<212> PRT

<213> Homo sapience

<400> 1

Met Ala Pro Gln Asn Leu Ser Thr Phe Cys Leu Leu Leu Tyr Leu

1 5 10 15

Ile Gly Ala Val Ile Ala Gly Arg Asp Phe Tyr Lys Ile Leu Gly Val

20 25 30

Pro Arg Ser Ala Ser Ile Lys Asp Ile Lys Lys Ala Tyr Arg Lys Leu

35 40 45

Ala Leu Gln Leu His Pro Asp Arg Asn Pro Asp Asp Pro Gln Ala Gln

50 55 60

Glu Lys Phe Gln Asp Leu Gly Ala Ala Tyr Glu Val Leu Ser Asp Ser

65 70 75 80

Glu Lys Arg Lys Gln Tyr Asp Thr Tyr Gly Glu Glu Gly Leu Lys Asp

.....

100 105 110

Phe Gly Phe Met Phe Gly Gly Thr Pro Arg Gln Gln Asp Arg Asn Ile			
115	120	125	
Pro Arg Gly Ser Asp Ile Ile Val Asp Leu Glu Val Thr Leu Glu Glu			
130	135	140	
Val Tyr Ala Gly Asn Phe Val Glu Val Val Arg Asn Lys Pro Val Ala			
145	150	155	160
Arg Gln Ala Pro Gly Lys Arg Lys Cys Asn Cys Arg Gln Glu Met Arg			
165	170	175	
Thr Thr Gln Leu Gly Pro Gly Arg Phe Gln Met Thr Gln Glu Val Val			
180	185	190	
Cys Asp Glu Cys Pro Asn Val Lys Leu Val Asn Glu Glu Arg Thr Leu			
195	200	205	
Glu Val Glu Ile Glu Pro Gly Val Arg Asp Gly Met Glu Tyr Pro Phe			
210	215	220	
Ile Gly Glu Gly Glu Pro His Val Asp Gly Glu Pro Gly Asp Leu Arg			
225	230	235	240
Phe Arg Ile Lys Val Val Lys His Pro Ile Phe Glu Arg Arg Gly Asp			
245	250	255	
Asp Leu Tyr Thr Asn Val Thr Ile Ser Leu Val Glu Ser Leu Val Gly			
260	265	270	
Phe Glu Met Asp Ile Thr His Leu Asp Gly His Lys Val His Ile Ser			
275	280	285	
Arg Asp Lys Ile Thr Arg Pro Gly Ala Lys Leu Trp Lys Lys Gly Glu			
290	295	300	
Gly Leu Pro Asn Phe Asp Asn Asn Asn Ile Lys Gly Ser Leu Ile Ile			
305	310	315	320
Thr Phe Asp Val Asp Phe Pro Lys Glu Gln Leu Thr Glu Glu Ala Arg			
325	330	335	
Glu Gly Ile Lys Gln Leu Leu Lys Gln Gly Ser Val Gln Lys Val Tyr			

340

345

350

Asn Gly Leu Gln Gly Tyr

355

【 0 0 6 4 】

<210> 2

<211> 226

<212> PRT

<213> Homo sapience

<400> 2

Met Lys Met Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys Cys

1

5

10

15

Leu Cys Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val Trp Tyr

20

25

30

Leu Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser Ala Leu Ala

35

40

45

Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu Gly Gly Asp Phe

50

55

60

Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala Ile Ala Ile Ser Leu

65

70

75

80

Leu Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly Ala Tyr Lys Gln

85

90

95

Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr Gln Ile Phe Asp Phe

100

105

110

Ala Leu Asn Met Leu Val Ala Ile Thr Val Leu Ile Tyr Pro Asn Ser

115

120

125

Ile Gln Glu Tyr Ile Arg Gln Leu Pro Pro Asn Phe Pro Tyr Arg Asp

Sp. var. 191.13. 191.13. 191.13. 191.13. 191.13. 191.13.

145

150

155

160

Phe Ile Ser Ile Ile Leu Thr Phe Lys Gly Tyr Leu Ile Ser Cys Val

165

170

175

Trp Asn Cys Tyr Arg Tyr Ile Asn Gly Arg Asn Ser Ser Asp Val Leu

180

185

190

Val Tyr Val Thr Ser Asn Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp

195

200

205

Asp Ala Thr Val Asn Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val

210

215

220

Ser Ala

225

【0 0 6 5】

<210> 3

<211> 195

<212> PRT

<213> Homo sapience

<400> 3

Met Arg Leu Leu Leu Leu Leu Leu Val Ala Ala Ser Ala Met Val Arg

1

5

10

15

Ser Glu Ala Ser Ala Asn Leu Gly Gly Val Pro Ser Lys Arg Leu Lys

20

25

30

Met Gln Tyr Ala Thr Gly Pro Leu Leu Lys Phe Gln Ile Cys Val Ser

35

40

45

Xaa Gly Tyr Arg Arg Val Phe Glu Glu Tyr Met Arg Val Ile Ser Gln

50

55

60

Arg Tyr Pro Asp Ile Arg Ile Glu Gly Glu Asn Tyr Leu Pro Gln Pro

65

70

75

80

Ile Tyr Arg His Ile Ala Ser Phe Leu Ser Val Phe Lys Leu Val Leu

85

90

95

Ile Gly Leu Ile Ile Val Gly Lys Asp Pro Phe Ala Phe Phe Gly Met



<211> 339

<213> Homo sapience

Met Asn Trp Glu Leu Leu Leu Trp Leu Leu Val Leu Cys Ala Leu Leu

III. $U(1)_A \times U(1)_B \times U(1)_C \times U(1)_D \times U(1)_E \times U(1)_F \times U(1)_G \times U(1)_H \times U(1)_I \times U(1)_J \times U(1)_K \times U(1)_L \times U(1)_M \times U(1)_N \times U(1)_O \times U(1)_P \times U(1)_Q \times U(1)_R \times U(1)_S \times U(1)_T \times U(1)_U \times U(1)_V \times U(1)_W \times U(1)_X \times U(1)_Y \times U(1)_Z \times U(1)_{\text{gravity}}$

Ala Arg Arg Val His Glu Leu Glu Arg Val Lys Arg Arg Cys Leu Glu			
85	90	95	
Asn Gly Asn Leu Lys Glu Lys Asp Ile Leu Val Leu Pro Leu Asp Leu			
100	105	110	
Thr Asp Thr Gly Ser His Glu Ala Ala Thr Lys Ala Val Leu Gln Glu			
115	120	125	
Phe Gly Arg Ile Asp Ile Leu Val Asn Asn Gly Gly Met Ser Gln Arg			
130	135	140	
Ser Leu Cys Met Asp Thr Ser Leu Asp Val Tyr Arg Lys Leu Ile Glu			
145	150	155	160
Leu Asn Tyr Leu Gly Thr Val Ser Leu Thr Lys Cys Val Leu Pro His			
165	170	175	
Met Ile Glu Arg Lys Gln Gly Lys Ile Val Thr Val Asn Ser Ile Leu			
180	185	190	
Gly Ile Ile Ser Val Pro Leu Ser Ile Gly Tyr Cys Ala Ser Lys His			
195	200	205	
Ala Leu Arg Gly Phe Phe Asn Gly Leu Arg Thr Glu Leu Ala Thr Tyr			
210	215	220	
Pro Gly Ile Ile Val Ser Asn Ile Cys Pro Gly Pro Val Gln Ser Asn			
225	230	235	240
Ile Val Glu Asn Ser Leu Ala Gly Glu Val Thr Lys Thr Ile Gly Asn			
245	250	255	
Asn Gly Asp Gln Ser His Lys Met Thr Thr Ser Arg Cys Val Arg Leu			
260	265	270	
Met Leu Ile Ser Met Ala Asn Asp Leu Lys Glu Val Trp Ile Ser Glu			
275	280	285	
Gln Pro Phe Leu Leu Val Thr Tyr Leu Trp Gln Tyr Met Pro Thr Trp			
290	295	300	
Ala Trp Trp Ile Thr Asn Lys Met Gly Lys Lys Arg Ile Glu Asn Phe			

305 310 315 320
 Lys Ser Gly Val Asp Ala Asp Ser Ser Tyr Phe Lys Ile Phe Lys Thr
 325 330 335
 Lys His Asp

【 0 0 6 7 】

<210> 5

<211> 487

<212> PRT

<213> Homo sapience

<400> 5

Met Asp Gly Thr Glu Thr Arg Gln Arg Arg Leu Asp Ser Cys Gly Lys
 1 5 10 15
 Pro Gly Glu Leu Gly Leu Pro His Pro Leu Ser Thr Gly Gly Leu Pro
 20 25 30
 Val Ala Ser Glu Asp Gly Ala Leu Arg Ala Pro Glu Ser Gln Ser Val
 35 40 45
 Thr Pro Lys Pro Leu Glu Thr Glu Pro Ser Arg Glu Thr Ala Trp Ser
 50 55 60
 Ile Gly Leu Gln Val Thr Val Pro Phe Met Phe Ala Gly Leu Gly Leu
 65 70 75 80
 Ser Trp Ala Gly Met Leu Leu Asp Tyr Phe Gln His Trp Pro Val Phe
 85 90 95
 Val Glu Val Lys Asp Leu Leu Thr Leu Val Pro Pro Leu Val Gly Leu
 100 105 110
 Lys Gly Asn Leu Glu Met Thr Leu Ala Ser Arg Leu Ser Thr Ala Ala

130

135

140

Ser Asn Leu Ala Leu Ile Gln Val Gln Ala Thr Val Val Gly Leu Leu
 145 150 155 160
 Ala Ala Val Ala Ala Leu Leu Leu Gly Val Val Ser Arg Glu Glu Val
 165 170 175
 Asp Val Ala Lys Val Glu Leu Leu Cys Ala Ser Ser Val Leu Thr Ala
 180 185 190
 Phe Leu Ala Ala Phe Ala Leu Gly Val Leu Met Val Cys Ile Val Ile
 195 200 205
 Gly Ala Arg Lys Leu Gly Val Asn Pro Asp Asn Ile Ala Thr Pro Ile
 210 215 220
 Ala Ala Ser Leu Gly Asp Leu Ile Thr Leu Ser Ile Leu Ala Leu Val
 225 230 235 240
 Ser Ser Phe Phe Tyr Arg His Lys Asp Ser Arg Tyr Leu Thr Pro Leu
 245 250 255
 Val Cys Leu Ser Phe Ala Ala Leu Thr Pro Val Trp Val Leu Ile Ala
 260 265 270
 Lys Gln Ser Pro Pro Ile Val Lys Ile Leu Lys Phe Gly Trp Phe Pro
 275 280 285
 Ile Ile Leu Ala Met Val Ile Ser Ser Phe Gly Gly Leu Ile Leu Ser
 290 295 300
 Lys Thr Val Ser Lys Gln Gln Tyr Lys Gly Met Ala Ile Phe Thr Pro
 305 310 315 320
 Val Ile Cys Gly Val Gly Gly Asn Leu Val Ala Ile Gln Thr Ser Arg
 325 330 335
 Ile Ser Thr Tyr Leu His Met Trp Ser Ala Pro Gly Val Leu Pro Leu
 340 345 350
 Gln Met Lys Lys Phe Trp Pro Asn Pro Cys Ser Thr Phe Cys Thr Ser
 355 360 365
 Glu Ile Asn Ser Met Ser Ala Arg Val Leu Leu Leu Leu Val Val Pro



<400> 6

$\Gamma_{\text{eff}} = \frac{\gamma}{\alpha} \left(\frac{1}{\beta} + \frac{\alpha}{\gamma} \right) \approx \frac{\gamma}{\alpha} \left(\frac{1}{\beta} + \frac{\alpha}{\gamma} \right)$

60

Arg	Ser	Tyr	Cys	Ser	Ala	Lys	Ala	Arg	Asp	Arg	His	Phe	Ala	Gly	Asp
65					70				75						80
Val	Leu	Gly	Tyr	Val	Thr	Pro	Trp	Asn	Ser	His	Gly	Tyr	Asp	Val	Thr
				85				90						95	
Lys	Val	Phe	Gly	Ser	Lys	Phe	Thr	Gln	Ile	Ser	Pro	Val	Trp	Leu	Gln
			100					105						110	
Leu	Lys	Arg	Arg	Gly	Arg	Glu	Met	Phe	Glu	Val	Thr	Gly	Leu	His	Asp
			115				120						125		
Val	Asp	Gln	Gly	Trp	Met	Arg	Ala	Val	Arg	Lys	His	Ala	Lys	Gly	Leu
	130					135							140		
His	Ile	Val	Pro	Arg	Leu	Leu	Phe	Glu	Asp	Trp	Thr	Tyr	Asp	Asp	Phe
145				150					155						160
Arg	Asn	Val	Leu	Asp	Ser	Glu	Asp	Glu	Ile	Glu	Glu	Leu	Ser	Lys	Thr
				165				170						175	
Val	Val	Gln	Val	Ala	Lys	Asn	Gln	His	Phe	Asp	Gly	Phe	Val	Val	Glu
			180					185						190	
Val	Trp	Asn	Gln	Leu	Leu	Ser	Gln	Lys	Arg	Val	Gly	Leu	Ile	His	Met
			195					200						205	
Leu	Thr	His	Leu	Ala	Glu	Ala	Leu	His	Gln	Ala	Arg	Leu	Leu	Ala	Leu
			210					215						220	
Leu	Val	Ile	Pro	Pro	Ala	Ile	Thr	Pro	Gly	Thr	Asp	Gln	Leu	Gly	Met
225					230					235					240
Phe	Thr	His	Lys	Glu	Phe	Glu	Gln	Leu	Ala	Pro	Val	Leu	Asp	Gly	Phe
				245					250					255	
Ser	Leu	Met	Thr	Tyr	Asp	Tyr	Ser	Thr	Ala	His	Gln	Pro	Gly	Pro	Asn
				260					265					270	
Ala	Pro	Leu	Ser	Trp	Val	Arg	Ala	Cys	Val	Gln	Val	Leu	Asp	Pro	Lys
			275					280						285	
Ser	Lys	Trp	Arg	Ser	Lys	Ile	Leu	Leu	Gly	Leu	Asn	Phe	Tyr	Gly	Met

290	295	300	
Asp Tyr Ala Thr Ser Lys Asp Ala Arg Glu Pro Val Val Gly Ala Arg			
305	310	315	320
Tyr Ile Gln Thr Leu Lys Asp His Arg Pro Arg Met Val Trp Asp Ser			
	325	330	335
Gln Ala Ser Glu His Phe Phe Glu Tyr Lys Lys Ser Arg Ser Gly Arg			
	340	345	350
His Val Val Phe Tyr Pro Thr Leu Lys Ser Leu Gln Val Arg Leu Glu			
	355	360	365
Leu Ala Arg Glu Leu Gly Val Gly Val Ser Ile Trp Glu Leu Gly Gln			
	370	375	380
Gly Leu Asp Tyr Phe Tyr Asp Leu Leu			
385	390		

【 0 0 6 9 】

<210> 7

<211> 196

<212> PRT

<213> Homo sapience

<400> 7

Met Trp Arg Val Pro Gly Thr Thr Arg Arg Pro Val Thr Gly Glu Ser			
1	5	10	15
Pro Gly Met His Arg Pro Glu Ala Met Leu Leu Leu Leu Thr Leu Ala			
	20	25	30
Leu Leu Gly Gly Pro Thr Trp Ala Gly Lys Met Tyr Gly Pro Gly Gly			
	35	40	45
Gly Lys Tyr Phe Ser Thr Thr Glu Asp Tyr Asp His Glu Ile Thr Gly			

eu Arg val ser val ty leu leu leu val lys ala val val val			
65	70	75	80

Leu Gly Asp Ser Trp Asp Val Lys Leu Gly Ala Leu Gly Gly Asn Thr

85

90

95

Gln Glu Val Thr Leu Gln Pro Gly Glu Tyr Ile Thr Lys Val Phe Val

100

105

110

Ala Phe Gln Ala Phe Leu Arg Gly Met Val Met Tyr Thr Ser Lys Asp

115

120

125

Arg Tyr Phe Tyr Phe Gly Lys Leu Asp Gly Gln Ile Ser Ser Ala Tyr

130

135

140

Pro Ser Gln Glu Gly Gln Val Leu Val Gly Ile Tyr Gly Gln Tyr Gln

145

150

155

160

Leu Leu Gly Ile Lys Ser Ile Gly Phe Glu Trp Asn Tyr Pro Leu Glu

165

170

175

Glu Pro Thr Thr Glu Pro Pro Val Asn Leu Thr Tyr Ser Ala Asn Ser

180

185

190

Pro Val Gly Arg

195

【 0 0 7 0 】

<210> 8

<211> 107

<212> PRT

<213> Homo sapience

<400> 8

Met Glu Gln Lys Leu Val Glu Glu Ile Leu Gln Ala Ile Thr Met Ser

1

5

10

15

Thr Asp Thr Gly Val Ser Leu Pro Ser Tyr Glu Glu Asp Gln Gly Ser

20

25

30

Lys Leu Ile Arg Lys Ala Lys Glu Ala Pro Phe Val Pro Val Gly Ile

35

40

45

Ala Gly Phe Ala Ala Ile Val Ala Tyr Gly Leu Tyr Lys Leu Lys Ser

50 55 60
 Arg Gly Asn Thr Lys Met Ser Ile His Leu Ile His Met Arg Val Ala
 65 70 75 80
 Ala Glu Gly Phe Val Val Gly Ala Met Thr Val Gly Met Gly Tyr Ser
 85 90 95
 Met Tyr Arg Glu Phe Trp Ala Lys Pro Lys Pro
 100 105

【 0 0 7 1 】

<210> 9

<211> 350

<212> PRT

<213> Homo sapience

<400> 9

Met Ser Glu Val Lys Ser Arg Lys Lys Ser Gly Pro Lys Gly Ala Pro
 1 5 10 15
 Ala Ala Glu Pro Gly Lys Arg Ser Glu Gly Gly Lys Thr Pro Val Ala
 20 25 30
 Arg Ser Ser Gly Gly Gly Gly Trp Ala Asp Pro Arg Thr Cys Leu Ser
 35 40 45
 Leu Leu Ser Leu Gly Thr Cys Leu Gly Leu Ala Trp Phe Val Phe Gln
 50 55 60
 Gln Ser Glu Lys Phe Ala Lys Val Glu Asn Gln Tyr Gln Leu Leu Lys
 65 70 75 80
 Leu Glu Thr Asn Glu Phe Gln Gln Leu Gln Ser Lys Ile Ser Leu Ile
 85 90 95
 Ser Glu Lys Trp Gln Lys Ser Glu Ala Ile Met Glu Gln Leu Lys Ser

115

120

125

Val Lys Thr Trp Ser Asn Arg Ile Thr Glu Lys Gln Asp Ile Leu Asn			
130	135	140	
Asn Ser Leu Thr Thr Leu Ser Gln Asp Ile Thr Lys Val Asp Gln Ser			
145	150	155	160
Thr Thr Ser Met Ala Lys Asp Val Gly Leu Lys Ile Thr Ser Val Lys			
165	170	175	
Thr Asp Ile Arg Arg Ile Ser Gly Leu Val Thr Asp Val Ile Ser Leu			
180	185	190	
Thr Asp Ser Val Gln Glu Leu Glu Asn Lys Ile Glu Lys Val Glu Lys			
195	200	205	
Asn Thr Val Lys Asn Ile Gly Asp Leu Leu Ser Ser Ser Ile Asp Arg			
210	215	220	
Thr Ala Thr Leu Arg Lys Thr Ala Ser Glu Asn Ser Gln Arg Ile Asn			
225	230	235	240
Ser Val Lys Lys Thr Leu Thr Glu Leu Lys Ser Asp Phe Asp Lys His			
245	250	255	
Thr Asp Arg Phe Leu Ser Leu Glu Gly Asp Arg Ala Lys Val Leu Lys			
260	265	270	
Thr Val Thr Phe Ala Asn Asp Leu Lys Pro Lys Val Tyr Asn Leu Lys			
275	280	285	
Lys Asp Phe Ser Arg Leu Glu Pro Leu Val Asn Asp Leu Thr Leu Arg			
290	295	300	
Ile Gly Arg Leu Val Thr Asp Leu Leu Gln Arg Glu Lys Glu Ile Ala			
305	310	315	320
Phe Leu Ser Glu Lys Ile Ser Asn Leu Thr Ile Val Gln Ala Glu Ile			
325	330	335	
Lys Asp Ile Lys Asp Glu Ile Ala His Ile Ser Asp Met Asn			
340	345	350	

【 0 0 7 2 】

<210> 10

<211> 107

<212> PRT

<213> Homo sapience

<400> 10

Met Ser Ser Ala Gly Thr Ala Thr Pro Leu Glu Met Asp His Lys Leu

1 5 10 15

Thr Ser Gln Pro Gly Arg Pro Ser Phe Tyr Cys Asn Ser Arg His Ser

20 25 30

Ile Val Gly Ser Ser His Gln Leu Gly Phe Trp Phe Ser His Leu Glu

35 40 45

Ser Ser Gly Leu Lys Val Phe Gln Val Ser Leu Pro Cys Glu Cys Val

50 55 60

Asn Leu Pro Thr Arg Ile Ala Ser Val Val Leu Ser Leu Met Ser Leu

65 70 75 80

Leu Val Val Gly Gln Ala Pro Ala Trp Glu Gly Ser Leu Leu Arg Gly

85 90 95

Arg Pro Ala Gly Gly Ala His Leu Cys Ala Ala

100 105

【 0 0 7 3 】

<210> 11

<211> 1074

<212> DNA

<213> Homo Sapience

<400> 11

atggctccgc agaacctgag caccttttgc ctgttgctgc tatacctcat cggggcgggtg 60

attaaaaagg cctataggaa actagccctg cagctccat cggccggaa cggcaga

ccacaagccc aggagaaatt ccaggatctg ggtgctgctt atgaggttct gtcagatagt 240

gagaaacgga aacagtacga tacttatggt gaagaaggat taaaagatgg tcatcagagc 300
 tcccatggag acattttttc acacttcttt ggggattttg gtttcatgtt tggaggaacc 360
 cctcgtcagc aagacagaaa tattccaaga ggaagtata ttattgtaga tctagaagtc 420
 actttggaag aagtatatgc aggaaatttt gtggaagtag ttagaaacaa acctgtggca 480
 aggcaggctc ctggcaaacg gaagtgcaat tgtcggcaag agatgcggac caccagctg 540
 ggccctgggc gcttccaaat gaccagagag gtggtctgcg acgaatgccc taatgtcaaa 600
 ctagtgaatg aagaacgaac gctggaagta gaaatagagc ctggggtgag agacggcatg 660
 gagtacccct ttattggaga aggtgagcct cacgtggatg gggagcctgg agatttacgg 720
 ttccgaatca aagttgtcaa gcaccaata ttgaaagga gaggagatga tttgtacaca 780
 aatgtgacaa tctcattagt tgagtcactg gttggctttg agatggatat tactcacttg 840
 gatggtcaca aggtacatat ttcccgggat aagatcacca ggccaggagc gaagctatgg 900
 aagaaagggg aagggtctcc caactttgac aacaacaata tcaagggtc tttgataatc 960
 acttttgatg tggattttcc aaaagaacag ttaacagagg aagcgagaga aggtatcaaa 1020
 cagctactga aacaagggtc agtgcagaag gtatacaatg gactgcaagg atat 1074

【 0 0 7 4 】

<210> 12

<211> 678

<212> DNA

<213> Homo Sapience

<400> 12

atgaagatgg tcgcgccttg gacgcggttc tactccaaca gctgctgctt gtgctgcat 60
 gtccgcaccg gcaccatcct gtcggcgctc tggatatga tcatcaatgc tgtggtactg 120
 ttgattttat tgagtgcctt ggctgatccg gatcagtata acttttcaag ttctgaactg 180
 ggaggtgact ttgagttcat ggatgatgcc aacatgtgca ttgccattgc gatttctctt 240
 ctcatgatcc tgatatgtgc tatggctact tacggagcgt acaagcaacg cgcagcctgg 300
 atcatcccat tcttctgtta ccagatcttt gactttgccc tgaacatgtt ggttgcaatc 360
 actgtgctta tttatccaaa ctccattcag gaatacatac ggcaactgcc tccatatttt 420
 ccctacagag atgatgtcat gtcagtgaat cctacctgtt tggtccttat tattcttctg 480
 tttattagca ttatcttgac ttttaagggt tacttgatta gctgtgtttg gaactgctac 540

cgatacatca atggtaggaa ctcctctgat gtcctggttt atgttaccag caatgacact 600
acgggtgctgc taccctcgta tgatgatgcc actgtgaatg gtgctgcaa ggagccaccg 660
ccaccttacg tgtctgcc 678

【0075】

<210> 13

<211> 585

<212> DNA

<213> Homo Sapience

<400> 13

atgaggcttc tgcctcttct cctagtggcg gcgtctgca tggccggag cgaggcctcg 60
gccaatctgg gcggcgtgcc cagcaagaga itaaagaigc agtacgccac ggggcccgtg 120
ctcaagticc agatttgtgt ttcctgaggt tataggcggg tgtttgagga gtacatgcgg 180
gttattagcc agcggtagcc agacatccgc attgaaggag agaattacct ccctcaacca 240
atatatagac acatagcacc ttccctgtca gtcttcaaac tagtattaat aggcttaata 300
attgttggca aggatccitt tgccttcttt ggcatgcaag ctccatagcat ctggcagtgg 360
ggccaagaaa ataagggtta tgcattgtat atggttttct tcttgagcaa catgattgag 420
aaccagtgtg tgtcaacagg tgcatttgag ataacttta atgatgtacc tgtgtggtct 480
aagctggaat ctggtagcct tccatccatg caacaacttg ttcaaattct tgacaatgaa 540
atgaagctca atgtgcatat ggattcaatc ccacaccatc gatca 585

【0076】

<210> 14

<211> 1017

<212> DNA

<213> Homo Sapience

<400> 14

atgaactggg agctgctgct gtggctgctg gtgctgtgcg cgctgctcct gctcttgggtg 60
ggacgaagcc cagaatggga gctgactgac atggctggctc gggtagctgg agccctggagt
ggaattgggtg aggagctggc ttaccagttg tctaaactag gagtttctct tgtgctgtca 240

gccagaagag tgcattgagct ggaaaggggtg aaaagaagat gcctagagaa tggcaattta 300
 aaagaaaaag atatacttgt ttigccccctt gacctgaccg acactgggttc ccatgaagcg 360
 gctaccaaag ctgttctcca ggagtttggt agaatcgaca ttctgggtcaa caatgggtgga 420
 atgtccacgc gttctctgtg catggatacc agcttggatg tctacagaaa gctaatagag 480
 ctttaactact tagggacggt gtccttgaca aaatgtgttc tgcctcacat gatcgagagg 540
 aagcaaggaa agattgttac tgtgaatagc atcctgggta tcatactctgt acctctttcc 600
 attggatact gtgctagcaa gcatgccttc cggggttttt ttaatggcct tcgaacagaa 660
 cttgccacat acccaggat aatagtttct aacatttgcc caggacctgt gcaatcaaatt 720
 attgtggaga attccctagc tggagaagtc acaaagacta taggcaataa tggagaccag 780
 tcccacaaga tgacaaccag tcgttgtgtg cggctgatgt taatcagcat ggccaatgat 840
 ttgaaagaag ttggatctc agaacaacct ttcttgtag taacatattt gtggcaatac 900
 atgccaacct gggcctgggtg gataaccaac aagatgggga agaaaaggat tgagaacttt 960
 aagagtgggtg tggatgcaga ctcttcttat tttaaaatct ttaagacaaa acatgac 1017

【 0 0 7 7 】

<210> 15

<211> 1461

<212> DNA

<213> Homo Sapience

<400> 15

atggatggga cagagacccg gcagcggagg ctggacagct gtggcaagcc aggggagctg 60
 gggcttctc accccctcag cacaggagga ctccctgtag cctcagaaga tggagctctc 120
 agggccccctg agagccaaag cgtgaccccc aagccactgg agactgagcc tagcaggagg 180
 accgcctggt ccataggcct tcaggtgacc gtgcccttca tgtttgcagg cctgggactg 240
 tcctgggccc gcatgcttct ggactatttc cagcactggc ctgtgtttgt ggaggtgaaa 300
 gaccttttga cattgggtgcc gcccctgggt ggcctgaagg ggaacctgga gatgacactg 360
 gcatccagac tctccacagc tgccaacact ggacaaattg atgaccccca ggagcagcac 420
 agagtcattc gcagcaacct ggccctcacc cagggtgcagg ccactgtcgt ggggctcttg 480
 gctgctgtgg ctgcgctgct gttggggcgt gtgtctcgag aggaagtgga tgtcgccaag 540
 gtggagtgtc tgtgtgccag cagtgtctc actgccttcc ttgcagcctt tgccctgggg 600

gtgctgatgg tctgtatagt gattgggtgct cgaaagctcg gggtaacccc agacaacatt	660
gccacgccc a ttgcagccag cctgggagac ctcacacac tgcattctt ggctttgggt	720
agcagcttct tctacagaca caaagatagt cggtaatcga cgccgctggt ctgcctcagc	780
tttgcggctc tgaccccagt gttgggtctc attgccaagc agagcccacc catcgtgaag	840
atcctgaagt ttggctgggt cccaatcacc ctggccatgg tcatcagcag ttctggagga	900
ctcatcttga gcaaaaccgt ttctaaccag cagtacaaag gcatggcgat atttaccccc	960
gtcatatgtg gtgttgggtg caatctgggt gccattcaga ccagccgaat ctcaacctac	1020
ctgcacatgt ggagtgcacc tggcgtcctg cccctccaga tgaagaaatt ctggcccaac	1080
ccgtgttcta ctttctgcac gtcagaaatc aattccatgt cagctcgagt cctgctcttg	1140
ctgggtgtcc caggccatct gattttcttc tacatcatct acctggtgga gggtcagtca	1200
gtcataaaca gccagacctt tgtgggtgct taccgtcgg caggccgat ccaggtgaca	1260
atcctgctgt acctggcaga agtgatgggt cggtgactt ggcaccaggc cctggatcct	1320
gacaaccact gcatcccta ccttacaggg ctgggggacc tgctcggtac tggcctcctg	1380
gcactctgct ttttactga ctggctactg aagagcaagg cagagctggg tggcatctca	1440
gaactggcat ctggacctcc c	1461

【 0 0 7 8 】

<210> 16

<211> 1179

<212> DNA

<213> Homo Sapience

<400> 16

atgcggacac tcttcaacct cctctggctt gccctggcct gcagccctgt tcacactacc	60
ctgtcaaagt cagatgccaa aaaagccgcc tcaaagacgc tgctggagaa gaggcagttt	120
tcagataagc cgggtgcaaga ccggggtttg gtggtgacgg acctcaaagc tgagagtgtg	180
gttcttgagc atgcagcta ctgctcggca aaggcccggg acagacactt tgctggggat	240
gtactgggct atgtcactcc atggaacagc catggctacg atgtcaccaa ggtctttggg	300

tttgagggtca cgggcttcca cgactgggac cagggtggg agcgagctgt cagggaagca	
gccaaagggcc tgcacatagt gctcggctc ctgtttgagg actggactta cgatgatttc	480

cggaacgtct tagacagtga ggatgagata gaggagctga gcaagaccgt ggtccagggtg 540
 gcaaagaacc agcatttcga tggcttcgtg gtggagggtct ggaaccagct gctaagccag 600
 aagcgcgtgg gccatcatca catgtcacc cacttggccg aggcctctgca ccaggcccgg 660
 ctgctggccc tccctggcat cccgcctgcc atcacccccg ggaccgacca gctgggcatg 720
 ttacgcaca aggagtittga gcagctggcc cccgtgctgg atggtttcag cctcatgacc 780
 tacgactact ctacagcgca tcagcctggc cctaattcac ccctgtcctg ggttcgagcc 840
 tgcgtccagg tccctggaccc gaagtccaag tggcgaagca aaatcctcct ggggctcaac 900
 ttctatggta tggactacgc gacctccaag gatgcccgtg agcctgttgt cggggccagg 960
 tacatccaga cactgaagga ccacaggccc cggatggtgt gggacagcca ggcctcagag 1020
 cacttcttcg agtacaagaa gagccgcagt gggaggcacg tcgtcttcta cccaacctg 1080
 aagtccttgc aggtgcggct ggagctggcc cgggagctgg gcgttgggggt ctctatctgg 1140
 gagctgggcc agggcctgga ctacttctac gacctgctc 1179

【0 0 7 9】

<210> 17

<211> 588

<212> DNA

<213> Homo Sapience

<400> 17

atgtggaggg tgcccggcac aaccagacgc ccagtcacag gcgagagccc tgggatgcac 60
 cggccagagg ccatgtctgt gctgctcacg cttgccctcc tggggggccc cacctgggca 120
 gggaagatgt atggccctgg aggaggcaag tatttcagca ccactgaaga ctacgaccat 180
 gaaatcacag ggctgcgggt gtctgtaggt cttctcctgg tgaaaagtgt ccagggtgaaa 240
 cttggagact cctgggacgt gaaactggga gccttaggtg ggaataccca ggaagtcacc 300
 ctgcagccag gcgaatacat cacaaaagtc tttgtgcct tccaagcttt cctccgggggt 360
 atggtcatgt acaccagcaa ggaccgctat ttctattttg ggaagcttga tggccagatc 420
 tccctgcct accccagcca agaggggagc gtgctgggtg gcatttatgg ccagtatcaa 480
 ctccctggca tcaagagcat tggctttgaa tgggaattatc cactagagga gccgaccact 540
 gagccaccag ttaatctcac atactcagca aactcaccg tgggtcgc 588

【0 0 8 0】

<210> 18

<211> 321

<212> DNA

<213> Homo Sapience

<400> 18

atggagcaga agcttgtgga ggagattcctt caagcaatca ctatgtcaac agacacaggt	60
gtttcccttc cttcatatga ggaagatcag ggatcaaaac tcattcgaaa agctaaagag	120
gcaccattcg taccctgttg aatagcgggt ttgcagcaa ttgttgata tggattatat	180
aaactgaaga gcaggggaaa tactaaaatg tccattcatt tgatccacat gcgtgtggca	240
gccaaggct ttgtttagg agcaatgact gtigtatgg gctattccat gtatcgggaa	300
ttctgggcaa aacctagcc t	321

【 0 0 8 1 】

<210> 19

<211> 1050

<212> DNA

<213> Homo Sapience

<400> 19

atgtctgagg tgaagagccg gaagaagtcg gggcccaagg gagccccctgc tgcggagccc	60
gggaagcggg gcgagggcgg gaagaccccc gtggcccggg gcagcggagg cgggggctgg	120
gcagaccccc gaacgtgcct gagcctgctg tcgctgggga cgtgcctggg cctggcctgg	180
tttgtatttc agcagtcaga aaaatttgca aaggctggaaa accaatacca gttactgaaa	240
ctagaaacca atgaattcca acaacttcaa agtaaaatca gtttaatttc agaaaagtgg	300
cagaaatctg aagctatcat ggaacaattg aagtcctttc aaataattgc tcatctaaag	360
cgtctacagg aagaaattaa tgaggtaaaa acttggtcca ataggataac tgaaaaacag	420
gatatactga acaacagtct gacgacgctt tctcaagaca ttacaaaagt agaccaaagt	480
acaacttcca tggcaaaaga tgttggtctc aagattacaa gtgtaaaaac agatatacga	540
atataaatag agtaagttag aaaaatttca gcttttttgc aggtgagc	600
agtattgatc gaacagcaac gctccgaag acagcattcg aaaattcaca aagaattaac	720

tctgttaaga agacgctaac cgaactaaag agtgacttcg acaaacatac agatagattt	780
ctaagcttag aaggtagacag agccaaagtt ctgaagacag tgacttttgc aaatgatcta	840
aaaccaaagg tgtataatct aaagaaggac ttttcccggt tagaaccatt agtaaagat	900
ttaacactac gcattgggag attggttacc gacttactac aaagagagaa agaaattgct	960
ttcttaagtg aaaaaatatac taatttaaca atagtccaag ctgagattaa ggatattaaa	1020
gatgaaatag cacacatttc agatatgaat	1050

【0 0 8 2】

<210> 20

<211> 321

<212> DNA

<213> Homo Sapience

<400> 20

atgtcctcag caggcacagc aacccctctg gaaatggatc acaaactcac ttctcagcca	60
ggcaggccaa gcttctattg taacagtagg cacagtatag tcggatcatt acatcagctg	120
ggtttttggg ttagtcatct agagtcgtct ggactaaagg tctttcaggt ctccctgccc	180
tgtgagtgcg tgaacctccc caccogaatt gcctcagttg tcttgagcct catgtctctc	240
ctggtgggtgg gccaggcccc tgcattgggaa gggagcctgc tgcggggcag gccagctggg	300
ggtgctcacc tatgcgcagc a	321

【0 0 8 3】

<210> 21

<211> 1619

<212> DNA

<213> Homo Sapience

<220>

<221> CDS

<222> (158)...(1234)

<400> 21

agaagagggg gctagctagc tgtctctgcg gaccaggag acccccgcg cccccgggtg	60
tgaggcggcc tcacagggcc gggtagggctg gcgagccgac gcggcggcgg aggaggctgt	120

gaggagtgtg tggaacagga cccgggacag aggaacc atg gct ccg cag aac ctg 175
Met Ala Pro Gln Asn Leu
1 5

agc acc ttt tgc ctg ttg ctg cta tac ctc atc ggg gcg gtg att gcc 223
Ser Thr Phe Cys Leu Leu Leu Leu Tyr Leu Ile Gly Ala Val Ile Ala
10 15 20

gga cga gat ttc tat aag atc ttg ggg gtg cct cga agt gcc tct ata 271
Gly Arg Asp Phe Tyr Lys Ile Leu Gly Val Pro Arg Ser Ala Ser Ile
25 30 35

aag gat att aaa aag gcc tat agg aaa cta gcc ctg cag ctt cat ccc 319
Lys Asp Ile Lys Lys Ala Tyr Arg Lys Leu Ala Leu Gln Leu His Pro
40 45 50

gac cgg aac cct gat gat cca caa gcc cag gag aaa ttc cag gat ctg 367
Asp Arg Asn Pro Asp Asp Pro Gln Ala Gln Glu Lys Phe Gln Asp Leu
55 60 65 70

ggt gct gct tat gag gtt ctg tca gat agt gag aaa cgg aaa cag tac 415
Gly Ala Ala Tyr Glu Val Leu Ser Asp Ser Glu Lys Arg Lys Gln Tyr
75 80 85

gat act tat ggt gaa gaa gga tta aaa gat ggt cat cag agc tcc cat 463
Asp Thr Tyr Gly Glu Glu Gly Leu Lys Asp Gly His Gln Ser Ser His
90 95 100

gga gac att ttt tca cac ttc ttt ggg gat ttt ggt ttc atg ttt gga 511
Gly Asp Ile Phe Ser His Phe Phe Gly Asp Phe Gly Phe Met Phe Gly
105 110 115

gga acc cct cgt cag caa gac aga aat att cca aga gga agt gat att 559
Gly Thr Pro Arg Gln Gln Asp Arg Asn Ile Pro Arg Gly Ser Asp Ile
120 125 130 135 140 145 150

Ile Val Asp Leu Glu Val Thr Leu Glu Glu Val Tyr Ala Gly Asn Phe

135	140	145	150	
gtg gaa gta gtt aga aac aaa cct gtg gca agg cag gct cct ggc aaa				655
Val Glu Val Val Arg Asn Lys Pro Val Ala Arg Gln Ala Pro Gly Lys				
155	160	165		
cgg aag tgc aat tgt cgg caa gag atg cgg acc acc cag ctg ggc cct				703
Arg Lys Cys Asn Cys Arg Gln Glu Met Arg Thr Thr Gln Leu Gly Pro				
170	175	180		
ggg cgc ttc caa atg acc cag gag gtg gtc tgc gac gaa tgc cct aat				751
Gly Arg Phe Gln Met Thr Gln Glu Val Val Cys Asp Glu Cys Pro Asn				
185	190	195		
gtc aaa cta gtg aat gaa gaa cga acg ctg gaa gta gaa ata gag cct				799
Val Lys Leu Val Asn Glu Glu Arg Thr Leu Glu Val Glu Ile Glu Pro				
200	205	210		
ggg gtg aga gac ggc atg gag tac ccc ttt att gga gaa ggt gag cct				847
Gly Val Arg Asp Gly Met Glu Tyr Pro Phe Ile Gly Glu Gly Glu Pro				
215	220	225	230	
cac gtg gat ggg gag cct gga gat tta cgg ttc cga atc aaa gtt gtc				895
His Val Asp Gly Glu Pro Gly Asp Leu Arg Phe Arg Ile Lys Val Val				
235	240	245		
aag cac cca ata ttt gaa agg aga gga gat gat ttg tac aca aat gtg				943
Lys His Pro Ile Phe Glu Arg Arg Gly Asp Asp Leu Tyr Thr Asn Val				
250	255	260		
aca atc tca tta gtt gag tca ctg gtt ggc ttt gag atg gat att act				991
Thr Ile Ser Leu Val Glu Ser Leu Val Gly Phe Glu Met Asp Ile Thr				
265	270	275		
cac ttg gat ggt cac aag gta cat att tcc cgg gat aag atc acc agg				1039
His Leu Asp Gly His Lys Val His Ile Ser Arg Asp Lys Ile Thr Arg				
280	285	290		
cca gga gcg aag cta tgg aag aaa ggg gaa ggg ctc ccc aac ttt gac				1087

Pro Gly Ala Lys Leu Trp Lys Lys Gly Glu Gly Leu Pro Asn Phe Asp	
295	300
305	310
aac aac aat atc aag ggc tct ttg ata atc act ttt gat gtg gat ttt	1135
Asn Asn Asn Ile Lys Gly Ser Leu Ile Ile Thr Phe Asp Val Asp Phe	
315	320
325	
cca aaa gaa cag tta aca gag gaa gcg aga gaa ggt atc aaa cag cta	1183
Pro Lys Glu Gln Leu Thr Glu Glu Ala Arg Glu Gly Ile Lys Gln Leu	
330	335
340	
ctg aaa caa ggg tca gtg cag aag gta tac aat gga ctg caa gga tat	1231
Leu Lys Gln Gly Ser Val Gln Lys Val Tyr Asn Gly Leu Gln Gly Tyr	
345	350
355	
tgagagtga ataaaattgg actttgttta aaataagtga ataagcgata tttattatct	1290
gcaagggtttt ttgtgtgtg tttttgtttt tattttcaat atgcaagtta ggcttaattt	1350
ttttatctaa tgatcatcat gaaatgaata agaggggctta agaatttgtc catttgcatt	1410
cggaaaagaa tgaccagcaa aagggtttact aataacctctc cctttgggga tttaatgtct	1470
ggtgctgccg cctgagtttc aagaattaaa gctgcaagag gactccagga gcaaaaagaaa	1530
cacaatatag agggttggag ttgttagcaa ttctattcaa aatgccaaact ggagaagtct	1590
gttttttaaat acattttgtt gttattttt	1619

【 0 0 8 4 】

<210> 22

<211> 2054

<212> DNA

<213> Homo Sapience

 $\langle 220 \rangle$

<221> CDS

<222> (254)...(934)

agcgaagggt accgaccggg cagaagctcg gagctctcgg ggtatcgagg aggcaggccc 120

gcggg'gcac gggcgagcgg gccgggagcc ggagcggcgg aggagccggc agcagcggcg	180
cggcgggctc caggcgaggc ggtcgacgct cctgaaaact tgcgcgcgcg ctcgcgccac	240
tgcgcccgga gcg atg aag atg gtc gcg ccc tgg acg cgg ttc tac tcc	289
Met Lys Met Val Ala Pro Trp Thr Arg Phe Tyr Ser	
1 5 10	
aac agc tgc tgc ttg tgc tgc cat gtc cgc acc ggc acc atc ctg ctc	337
Asn Ser Cys Cys Leu Cys Cys His Val Arg Thr Gly Thr Ile Leu Leu	
15 20 25	
ggc gtc tgg tat ctg atc atc aat gct gtg gta ctg ttg att tta ttg	385
Gly Val Trp Tyr Leu Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu	
30 35 40	
agt gcc ctg gct gat ccg gat cag tat aac ttt tca agt tct gaa ctg	433
Ser Ala Leu Ala Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu	
45 50 55 60	
gga ggt gac ttt gag ttc atg gat gat gcc aac atg tgc att gcc att	481
Gly Gly Asp Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala Ile	
65 70 75	
gcg att tct ctt ctc atg atc ctg ata tgt gct atg gct act tac gga	529
Ala Ile Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly	
80 85 90	
gcg tac aag caa cgc gca gcc tgg atc atc cca ttc ttc tgt tac cag	577
Ala Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr Gln	
95 100 105	
atc ttt gac ttt gcc ctg aac atg ttg gtt gca atc act gtg ctt att	625
Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile Thr Val Leu Ile	
110 115 120	
tat cca aac tcc att cag gaa tac ata cgg caa ctg cct cct aat ttt	673
Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln Leu Pro Pro Asn Phe	
125 130 135 140	

ccc tac aga gat gat gtc atg tca gtg aat cct acc tgt ttg gtc ctt	721
Pro Tyr Arg Asp Asp Val Met Ser Val Asn Pro Thr Cys Leu Val Leu	
145 150 155	
att att ctt ctg ttt att agc att atc ttg act ttt aag ggt tac ttg	769
Ile Ile Leu Leu Phe Ile Ser Ile Ile Leu Thr Phe Lys Gly Tyr Leu	
160 165 170	
att agc tgt gtt tgg aac tgc tac cga tac atc aat ggt agg aac tcc	817
Ile Ser Cys Val Trp Asn Cys Tyr Arg Tyr Ile Asn Gly Arg Asn Ser	
175 180 185	
tct gat gtc ctg gtt tat gtt acc agc aat gac act acg gtg ctg cta	865
Ser Asp Val Leu Val Tyr Val Thr Ser Asn Asp Thr Thr Val Leu Leu	
190 195 200	
ccc ccg tat gat gat gcc act gtg aat ggt gct gcc aag gag cca ccg	913
Pro Pro Tyr Asp Asp Ala Thr Val Asn Gly Ala Ala Lys Glu Pro Pro	
205 210 215 220	
cca cct tac gtg tct gcc taagccttca agtgggcgga gctgagggc	960
Pro Pro Tyr Val Ser Ala	
225	
agcagcttga ctttgcagac atctgagcaa tagttctgtt atttcacttt tgccatgagc	1020
ctctctgagc ttgtttgttg ctgaaatgct actttttaaa atttagatgt tagattgaaa	1080
actgtagttt tcaacatatg ctttgcctgga acactgtgat agattaactg tagaattctt	1140
cctgtacgat tggggatata atgggcttca ctaaccttcc ctaggcattg aaacttcccc	1200
caaactctgat ggacctagaa gtctgtcttt gtacctgtg ggccecaaag ttgggcattt	1260
ttctctctgt tccctctctt ttgaaaatgt aaaataaaac caaaaataga caactttttc	1320
ttcagccatt ccagcataga gaacaaaacc ttatggaaac aggaatgtca atttgttaat	1380
catgtttcta attaggtaaa tagaagtcct tatgtatgtg ttacaagaat ttccccaca	1440
tggectgdat taagaccatt agaaagcatt agggcgtggg gctagctgac cctgctgac	
tgttcttgtg gatcttgtgt ccagggacat ggggtgacat gcctcgtatg tgtttagaggg	1620

tggaatggat gtgtttggcg ctgcatggga tctgggtgcc ctcttctcct ggattcacat 1680
 cccacccag ggcccgcctt tactaagtgt tctgccctag attggttcaa ggaggtcatc 1740
 caactgactt tatcaagtgg aattgggata tatttgatat acttctgcct aacaacatgg 1800
 aaaagggttt tcttttcctt gcaagctaca tcctactgct ttgaacttcc aagtatgtct 1860
 agtcaccttt taaaatgtaa acattttcag aaaaatgagg attgccttcc ttgtatgcgc 1920
 tttttacctt gactacctga attgcaaggg atttttatat attcatatgt tacaaagtca 1980
 gcaactctcc tgttggttca ttattgaatg tgctgtaaat taagttgttt gcaattaaaa 2040
 caaggtttgc ccac 2054

【 0 0 8 5 】

<210> 23

<211> 1380

<212> DNA

<213> Homo Sapience

<220>

<221> CDS

<222> (43)...(630)

<400> 23

gcagtctgtc tgagggcggc cgaagtggct ggctcattta ag atg agg ctt ctg 54

Met Arg Leu Leu

1

ctg ctt ctc cta gtg gcg gcg tct gcg atg gtc cgg agc gag gcc tcg 102

Leu Leu Leu Leu Val Ala Ala Ser Ala Met Val Arg Ser Glu Ala Ser

5

10

15

20

gcc aat ctg ggc ggc gtg ccc agc aag aga tta aag atg cag tac gcc 150

Ala Asn Leu Gly Gly Val Pro Ser Lys Arg Leu Lys Met Gln Tyr Ala

25

30

35

acg ggg ccg ctg ctc aag ttc cag att tgt gtt tcc tga ggt tat agg 198

Thr Gly Pro Leu Leu Lys Phe Gln Ile Cys Val Ser Xaa Gly Tyr Arg

40

45

50

cgg gtg ttt gag gag tac atg cgg gtt att agc cag cgg tac cca gac	246
Arg Val Phe Glu Glu Tyr Met Arg Val Ile Ser Gln Arg Tyr Pro Asp	
55 60 65	
atc cgc att gaa gga gag aat tac ctc cct caa cca ata tat aga cac	294
Ile Arg Ile Glu Gly Glu Asn Tyr Leu Pro Gln Pro Ile Tyr Arg His	
70 75 80	
ata gca tct ttc ctg tca gtc ttc aaa cta gta tta ata ggc tta ata	342
Ile Ala Ser Phe Leu Ser Val Phe Lys Leu Val Leu Ile Gly Leu Ile	
85 90 95 100	
att gtt ggc aag gat cct ttt gct ttc ttt ggc atg caa gct cct agc	390
Ile Val Gly Lys Asp Pro Phe Ala Phe Phe Gly Met Gln Ala Pro Ser	
105 110 115	
atc tgg cag tgg ggc caa gaa aat aag gtt tat gca tgt atg atg gtt	438
Ile Trp Gln Trp Gly Gln Glu Asn Lys Val Tyr Ala Cys Met Met Val	
120 125 130	
ttc ttc ttg agc aac atg att gag aac cag tgt atg tca aca ggt gca	486
Phe Phe Leu Ser Asn Met Ile Glu Asn Gln Cys Met Ser Thr Gly Ala	
135 140 145	
ttt gag ata act tta aat gat gta cct gtg tgg tct aag ctg gaa tct	534
Phe Glu Ile Thr Leu Asn Asp Val Pro Val Trp Ser Lys Leu Glu Ser	
150 155 160	
ggc cac ctt cca tcc atg caa caa ctt gtt caa att ctt gac aat gaa	582
Gly His Leu Pro Ser Met Gln Gln Leu Val Gln Ile Leu Asp Asn Glu	
165 170 175 180	
atg aag ctc aat gtg cat atg gat tca atc cca cac cat cga tca	627
Met Lys Leu Asn Val His Met Asp Ser Ile Pro His His Arg Ser	
tag caccacctat cagcactgaa aactctctg catcaggga caacggcagg	
agcagcgtga ctgacattat gaaggcctgt actgaagaca gcaagctgtt agtacagacc	740

agatgctttc ttggcaggct cgttgtacct ctggaaaac ctcaatgcaa gatagtgttt 800
 cagtgcctggc atattttgga attctgcaca ttcatggagt gcaataatac tgtatagctt 860
 tccccacctc ccacaaaatc acccagttaa tgtgtgtgtg tgtttttttt ttttaaggtaa 920
 acattactac ttgtaacttt ttttcttagt catatttgaa aaagtagaaa attgagttac 980
 aatttgattt tttttccaaa gatgtctgtt aaatctgttg tgcttttata tgaatatttg 1040
 ttttttatag tttaaaattg atcctttggg aatccagttg aagttcccaa atactttata 1100
 agagtttatc agacatctct aatttggcca tgtccagttt atacagttta caaaatatag 1160
 cagatgcaag attatggggg aaatcctata ttcagagtac tctataaatt tttgtgtatg 1220
 tgtgtatgtg cgtgtgatta ccagagaact actaaaaaaaa ccaactgctt tttaaatcct 1280
 attgtgtagt taaagtgtca tgccttgacc aatctaataa attgattaat taactgggcc 1340
 tttatactta actaaataaa aaactaagca gatatgagtt 1380

【 0 0 8 6 】

<210> 24

<211> 1292

<212> DNA

<213> Homo Sapience

<220>

<221> CDS

<222> (113)...(1132)

<400> 24

aaaagtgcgg ctctgggctg gccgaagggg tggcgtgcg atcccgcagg gcagcgacgc 60
 gactctggtg cgggccgtct tcttcccccc gagctgggcg tgcgcggccg ca atg aac 118
 Met Asn

1

tgg gag ctg ctg ctg tgg ctg ctg gtg ctg tgc gcg ctg ctc ctg ctc 166
 Trp Glu Leu Leu Leu Trp Leu Leu Val Leu Cys Ala Leu Leu Leu Leu
 5 10 15
 ttg gtg cag ctg ctg cgc ttc ctg agg gct gac ggc gac ctg acg cta 214
 Leu Val Gln Leu Leu Arg Phe Leu Arg Ala Asp Gly Asp Leu Thr Leu

20	25	30	
cta tgg gcc gag tgg cag gga cga cgc cca gaa tgg gag ctg act gat			262
Leu Trp Ala Glu Trp Gln Gly Arg Arg Pro Glu Trp Glu Leu Thr Asp			
35	40	45	50
atg gtg gtg tgg gtg act gga gcc tcg agt gga att ggt gag gag ctg			310
Met Val Val Trp Val Thr Gly Ala Ser Ser Gly Ile Gly Glu Glu Leu			
	55	60	65
gct tac cag ttg tct aaa cta gga gtt tct ctt gtg ctg tca gcc aga			358
Ala Tyr Gln Leu Ser Lys Leu Gly Val Ser Leu Val Leu Ser Ala Arg			
	70	75	80
aga gtg cat gag ctg gaa agg gtg aaa aga aga tgc cta gag aat ggc			406
Arg Val His Glu Leu Glu Arg Val Lys Arg Arg Cys Leu Glu Asn Gly			
	85	90	95
aat tta aaa gaa aaa gat ata ctt gtt ttg ccc ctt gac ctg acc gac			454
Asn Leu Lys Glu Lys Asp Ile Leu Val Leu Pro Leu Asp Leu Thr Asp			
100	105	110	
act ggt tcc cat gaa gcg gct acc aaa gct gtt ctc cag gag ttt ggt			502
Thr Gly Ser His Glu Ala Ala Thr Lys Ala Val Leu Gln Glu Phe Gly			
115	120	125	130
aga atc gac att ctg gtc aac aat ggt gga atg tcc cag cgt tct ctg			550
Arg Ile Asp Ile Leu Val Asn Asn Gly Gly Met Ser Gln Arg Ser Leu			
	135	140	145
tgc atg gat acc agc ttg gat gtc tac aga aag cta ata gag ctt aac			598
Cys Met Asp Thr Ser Leu Asp Val Tyr Arg Lys Leu Ile Glu Leu Asn			
	150	155	160
tac tta ggg acg gtg tcc ttg aca aaa tgt gtt ctg cct cac atg atc			646
gag agg aag caa gga aag att gtt act gtg aat agc atc ctg ggt atc			694

Glu Arg Lys Gln Gly Lys Ile Val Thr Val Asn Ser Ile Leu Gly Ile	
180	185
ata tct gta cct ctt tcc att gga tac tgt gct agc aag cat gct ctc	742
Ile Ser Val Pro Leu Ser Ile Gly Tyr Cys Ala Ser Lys His Ala Leu	
195	200
cgg ggt ttt ttt aat ggc ctt cga aca gaa ctt gcc aca tac cca ggt	790
Arg Gly Phe Phe Asn Gly Leu Arg Thr Glu Leu Ala Thr Tyr Pro Gly	
215	220
ata ata gtt tct aac att tgc cca gga cct gtg caa tca aat att gtg	838
Ile Ile Val Ser Asn Ile Cys Pro Gly Pro Val Gln Ser Asn Ile Val	
230	235
gag aat tcc cta gct gga gaa gtc aca aag act ata ggc aat aat gga	886
Glu Asn Ser Leu Ala Gly Glu Val Thr Lys Thr Ile Gly Asn Asn Gly	
245	250
gac cag tcc cac aag atg aca acc agt cgt tgt gtg cgg ctg atg tta	934
Asp Gln Ser His Lys Met Thr Thr Ser Arg Cys Val Arg Leu Met Leu	
260	265
atc agc atg gcc aat gat ttg aaa gaa gtt tgg atc tca gaa caa cct	982
Ile Ser Met Ala Asn Asp Leu Lys Glu Val Trp Ile Ser Glu Gln Pro	
275	280
ttc ttg tta gta aca tat ttg tgg caa tac atg cca acc tgg gcc tgg	1030
Phe Leu Leu Val Thr Tyr Leu Trp Gln Tyr Met Pro Thr Trp Ala Trp	
295	300
tgg ata acc aac aag atg ggg aag aaa agg att gag aac ttt aag agt	1078
Trp Ile Thr Asn Lys Met Gly Lys Lys Arg Ile Glu Asn Phe Lys Ser	
310	315
ggt gtg gat gca gac tct tct tat ttt aaa atc ttt aag aca aaa cat	1126
Gly Val Asp Ala Asp Ser Ser Tyr Phe Lys Ile Phe Lys Thr Lys His	
325	330
	335

Asp

1240

1292

<210> 25

<211> 2168

<212> DNA

<213> Homo Sapience

 $\langle 220 \rangle$

<221> CDS

<222> (56)...(1519)

<400> 25

55

103

Met Asp Gly Thr Glu Thr Arg Gln Arg Arg Leu Asp Ser Cys Gly Lys

1 5 10 15

151

Pro Gly Glu Leu Gly Leu Pro His Pro Leu Ser Thr Gly Gly Leu Pro

20 25 30

199

Val Ala Ser Glu Asp Gly Ala Leu Arg Ala Pro Glu Ser Gln Ser Val

35 40 45

247

Thr Pro Lys Pro Leu Glu Thr Glu Pro Ser Arg Glu Thr Ala Trp Ser

am 22. Tag des Monats April 1900

Ile Gly Leu Gln Val Thr Val Pro Phe Met Phe Ala Gly Leu Gly Leu

65	70	75	80	
tcc tgg gcc ggc atg ctt ctg gac tat ttc cag cac tgg cct gtg ttt				343
Ser Trp Ala Gly Met Leu Leu Asp Tyr Phe Gln His Trp Pro Val Phe				
	85	90	95	
gtg gag gtg aaa gac ctt ttg aca ttg gtg ccg ccc ctg gtg ggc ctg				391
Val Glu Val Lys Asp Leu Leu Thr Leu Val Pro Pro Leu Val Gly Leu				
	100	105	110	
aag ggg aac ctg gag atg aca ctg gca tcc aga ctc tcc aca gct gcc				439
Lys Gly Asn Leu Glu Met Thr Leu Ala Ser Arg Leu Ser Thr Ala Ala				
	115	120	125	
aac act gga caa att gat gac ccc cag gag cag cac aga gtc atc agc				487
Asn Thr Gly Gln Ile Asp Asp Pro Gln Glu Gln His Arg Val Ile Ser				
	130	135	140	
agc aac ctg gcc ctc atc cag gtg cag gcc act gtc gtg ggg ctc ttg				535
Ser Asn Leu Ala Leu Ile Gln Val Gln Ala Thr Val Val Gly Leu Leu				
	145	150	155	160
gct gct gtg gct gcg ctg ctg ttg ggc gtg gtg tct cga gag gaa gtg				583
Ala Ala Val Ala Ala Leu Leu Leu Gly Val Val Ser Arg Glu Glu Val				
	165	170	175	
gat gtc gcc aag gtg gag ttg ctg tgt gcc agc agt gtc ctc act gcc				631
Asp Val Ala Lys Val Glu Leu Leu Cys Ala Ser Ser Val Leu Thr Ala				
	180	185	190	
ttc ctt gca gcc ttt gcc ctg ggg gtg ctg atg gtc tgt ata gtg att				679
Phe Leu Ala Ala Phe Ala Leu Gly Val Leu Met Val Cys Ile Val Ile				
	195	200	205	
ggt gct cga aag ctc ggg gtc aac cca gac aac att gcc acg ccc att				727
Gly Ala Arg Lys Leu Gly Val Asn Pro Asp Asn Ile Ala Thr Pro Ile				
	210	215	220	
gca gcc agc ctg gga gac ctc atc aca ctg tcc att ctg gct ttg gtt				775

Ala	Ala	Ser	Leu	Gly	Asp	Leu	Ile	Thr	Leu	Ser	Ile	Leu	Ala	Leu	Val	
225					230					235					240	
agc	agc	ttc	ttc	tac	aga	cac	aaa	gat	agt	cgg	tat	ctg	acg	ccg	ctg	823
Ser	Ser	Phe	Phe	Tyr	Arg	His	Lys	Asp	Ser	Arg	Tyr	Leu	Thr	Pro	Leu	
				245					250						255	
gtc	tgc	ctc	agc	ttt	gcg	gct	ctg	acc	cca	gtg	tgg	gtc	ctc	att	gcc	871
Val	Cys	Leu	Ser	Phe	Ala	Ala	Leu	Thr	Pro	Val	Trp	Val	Leu	Ile	Ala	
				260					265						270	
aag	cag	agc	cca	ccc	atc	gtg	aag	atc	ctg	aag	ttt	ggc	tgg	ttc	cca	919
Lys	Gln	Ser	Pro	Pro	Ile	Val	Lys	Ile	Leu	Lys	Phe	Gly	Trp	Phe	Pro	
				275					280						285	
atc	atc	ctg	gcc	atg	gtc	atc	agc	agt	ttc	gga	gga	ctc	atc	ttg	agc	967
Ile	Ile	Leu	Ala	Met	Val	Ile	Ser	Ser	Phe	Gly	Gly	Leu	Ile	Leu	Ser	
				290					295						300	
aaa	acc	gtt	tct	aaa	cag	cag	tac	aaa	ggc	atg	gcg	ata	ttt	acc	ccc	1015
Lys	Thr	Val	Ser	Lys	Gln	Gln	Tyr	Lys	Gly	Met	Ala	Ile	Phe	Thr	Pro	
305					310					315					320	
gtc	ata	tgt	ggt	gtt	ggt	ggc	aat	ctg	gtg	gcc	att	cag	acc	agc	cga	1063
Val	Ile	Cys	Gly	Val	Gly	Gly	Asn	Leu	Val	Ala	Ile	Gln	Thr	Ser	Arg	
				325					330						335	
atc	tca	acc	tac	ctg	cac	atg	tgg	agt	gca	cct	ggc	gtc	ctg	ccc	ctc	1111
Ile	Ser	Thr	Tyr	Leu	His	Met	Trp	Ser	Ala	Pro	Gly	Val	Leu	Pro	Leu	
				340					345						350	
cag	atg	aag	aaa	ttc	tgg	ccc	aac	ccg	tgt	tct	act	ttc	tgc	acg	tca	1159
Gln	Met	Lys	Lys	Phe	Trp	Pro	Asn	Pro	Cys	Ser	Thr	Phe	Cys	Thr	Ser	
				355					360						365	

© 2000 Blackwell Science Ltd *Journal of Internal Medicine* 247: 389–395

370 375 380

ggc cat ctg att ttc ttc tac atc atc tac ctg gtg gag ggt cag tca	1255
Gly His Leu Ile Phe Phe Tyr Ile Ile Tyr Leu Val Glu Gly Gln Ser	
385 390 395 400	
gtc ata aac agc cag acc ttt gtg gtg ctc tac ctg ctg gca ggc ctg	1303
Val Ile Asn Ser Gln Thr Phe Val Val Leu Tyr Leu Leu Ala Gly Leu	
405 410 415	
atc cag gtg aca atc ctg ctg tac ctg gca gaa gtg atg gtt cgg ctg	1351
Ile Gln Val Thr Ile Leu Leu Tyr Leu Ala Glu Val Met Val Arg Leu	
420 425 430	
act tgg cac cag gcc ctg gat cct gac aac cac tgc atc ccc tac ctt	1399
Thr Trp His Gln Ala Leu Asp Pro Asp Asn His Cys Ile Pro Tyr Leu	
435 440 445	
aca ggg ctg ggg gac ctg ctc ggt act ggc ctc ctg gca ctc tgc ttt	1447
Thr Gly Leu Gly Asp Leu Leu Gly Thr Gly Leu Leu Ala Leu Cys Phe	
450 455 460	
ttc act gac tgg cta ctg aag agc aag gca gag ctg ggt ggc atc tca	1495
Phe Thr Asp Trp Leu Leu Lys Ser Lys Ala Glu Leu Gly Gly Ile Ser	
465 470 475 480	
gaa ctg gca tct gga cct ccc taactgggcc ccgtgtgtcc catttgctca ttag	1550
Glu Leu Ala Ser Gly Pro Pro	
485	
aatttccctct cacatcagtg ggatacagaa ttcagtttct cccctgccag gtccttggga	1610
tggttgaccc ctgcctctgc agtagccttt tgtgagtctg ctaaggtagc tctcacacac	1670
ctcggtctctg gggttgatac ctgagcctgc aatagagccc tgaaatcaag agcatggctt	1730
gagtgtgtga atatgatgtg tgcacatgct taatgagcgt gcaagtgtgc acacgtttgt	1790
ggagaggagg gtgttctggc ctgagaagct aaagaagagg catgtccagt atgctttgca	1850
gggtgtgttt gctcttttcc atgcccattgc aaccagatt ggggtggagc aggaaggagc	1910
tcttttctgt tcccaagcct cagaactctt gagctgtggc ttacttgctg tcttcaccag	1970
gttcaagctc cgtgggccac actgtgtgtg tgccaagaag gtgtacagcc tccccaggat	2030

ggggcctcat acaacccttc atctgcactc aacatttaac cgtgtccttg ctgtcttttt 2090
 attttccttt ttgttagcaa aaacctctat ttagatttca ataatacagag aagtgtaaaa 2150
 taaaacagat tatattgt 2168

[0 0 8 8]

<210> 26

<211> 1357

<212> DNA

<213> Homo Sapience

<220>

<221> CDS

<222> (81)...(1262)

<400> 26

cgtgcgtttg tggccgtccg gcctccctga catgcagccc tctggacccc gaggttggac 60
 cctactgtga cacacctacc atg cgg aca ctc ttc aac ctc ctc tgg ctt 110

Met Arg Thr Leu Phe Asn Leu Leu Trp Leu

1 5 10

gcc ctg gcc tgc agc cct gtt cac act acc ctg tca aag tca gat gcc 158
 Ala Leu Ala Cys Ser Pro Val His Thr Thr Leu Ser Lys Ser Asp Ala

15 20 25

aaa aaa gcc gcc tca aag acg ctg ctg gag aag agt cag ttt tca gat 203
 Lys Lys Ala Ala Ser Lys Thr Leu Leu Glu Lys Ser Gln Phe Ser Asp

30 35 40

aag ccg gtg caa gac cgg ggt ttg gtg gtg acg gac ctc aaa gct gag 254
 Lys Pro Val Gln Asp Arg Gly Leu Val Val Thr Asp Leu Lys Ala Glu

45 50 55

agt gtg gtt ctt gag cat cgc agc tac tgc tcg gca aag gcc cgg gac 302

aga cac ttt gct ggg gat gta ctg ggc tat gtc act cca tgg aac agc 350



acc gac cag ctg ggc atg ttc acg cac aag gag ttt gag cag ctg gcc	830
Thr Asp Gln Leu Gly Met Phe Thr His Lys Glu Phe Glu Gln Leu Ala	
235 240 245 250	
ccc gtg ctg gat ggt ttc agc ctc atg acc tac gac tac tct aca gcg	878
Pro Val Leu Asp Gly Phe Ser Leu Met Thr Tyr Asp Tyr Ser Thr Ala	
255 260 265	
cat cag cct ggc cct aat gca ccc ctg tcc tgg gtt cga gcc tgc gtc	926
His Gln Pro Gly Pro Asn Ala Pro Leu Ser Trp Val Arg Ala Cys Val	
270 275 280	
cag gtc ctg gac ccg aag tcc aag tgg cga agc aaa atc ctc ctg ggg	974
Gln Val Leu Asp Pro Lys Ser Lys Trp Arg Ser Lys Ile Leu Leu Gly	
285 290 295	
ctc aac ttc tat ggt atg gac tac gcg acc tcc aag gat gcc cgt gag	1022
Leu Asn Phe Tyr Gly Met Asp Tyr Ala Thr Ser Lys Asp Ala Arg Glu	
300 305 310	
cct gtt gtc ggg gcc agg tac atc cag aca ctg aag gac cac agg ccc	1070
Pro Val Val Gly Ala Arg Tyr Ile Gln Thr Leu Lys Asp His Arg Pro	
315 320 325 330	
cgg atg gtg tgg gac agc cag gcc tca gag cac ttc ttc gag tac aag	1118
Arg Met Val Trp Asp Ser Gln Ala Ser Glu His Phe Phe Glu Tyr Lys	
335 340 345	
aag agc cgc agt ggg agg cac gtc gtc ttc tac cca acc ctg aag tcc	1166
Lys Ser Arg Ser Gly Arg His Val Val Phe Tyr Pro Thr Leu Lys Ser	
350 355 360	
ctg cag gtg cgg ctg gag ctg gcc cgg gag ctg ggc gtt ggg gtc tct	1214
Leu Gln Val Arg Leu Glu Leu Ala Arg Glu Leu Gly Val Gly Val Ser	
atc tgg ggc ctg ggc cag ggc ctg ggc cag ggc cag ggc cag ggc	
Ile Trp Glu Leu Gly Gln Gly Leu Asp Tyr Phe Tyr Asp Leu Leu	

380	385	390	
agggtgggcat tgcggccctcc gcggtaggacg tgttcctttc taagccatgg agtgagtgag			1320
cagggtgtgaa atacaggcct ccactccgtt tgctgtg			1357
【 0 0 8 9 】			
<210> 27			
<211> 711			
<212> DNA			
<213> Homo Sapience			
<220>			
<221> CDS			
<222> (8)...(598)			
<400> 27			
aaaggcg atg tgg agg gtg ccc ggc aca acc aga cgc cca gtc aca ggc			49
Met Trp Arg Val Pro Gly Thr Thr Arg Arg Pro Val Thr Gly			
1	5	10	
gag agc cct ggg atg cac cgg cca gag gcc atg ctg ctg ctg ctc acg			97
Glu Ser Pro Gly Met His Arg Pro Glu Ala Met Leu Leu Leu Leu Thr			
15	20	25	30
ctt gcc ctc ctg ggg ggc ccc acc tgg gca ggg aag atg tat ggc cct			145
Leu Ala Leu Leu Gly Gly Pro Thr Trp Ala Gly Lys Met Tyr Gly Pro			
35	40	45	
gga gga ggc aag tat ttc agc acc act gaa gac tac gac cat gaa atc			193
Gly Gly Gly Lys Tyr Phe Ser Thr Thr Glu Asp Tyr Asp His Glu Ile			
50	55	60	
aca ggg ctg cgg gtg tct gta ggt ctt ctc ctg gtg aaa agt gtc cag			241
Thr Gly Leu Arg Val Ser Val Gly Leu Leu Leu Val Lys Ser Val Gln			
65	70	75	
gtg aaa ctt gga gac tcc tgg gac gtg aaa ctg gga gcc tta ggt ggg			289
Val Lys Leu Gly Asp Ser Trp Asp Val Lys Leu Gly Ala Leu Gly Gly			

80	85	90	
aat acc cag gaa gtc acc ctg cag cca ggc gaa tac atc aca aaa gtc			337
Asn Thr Gln Glu Val Thr Leu Gln Pro Gly Glu Tyr Ile Thr Lys Val			
95	100	105	110
ttt gtc gcc ttc caa gct ttc ctc cgg ggt atg gtc atg tac acc agc			385
Phe Val Ala Phe Gln Ala Phe Leu Arg Gly Met Val Met Tyr Thr Ser			
115	120	125	
aag gac cgc tat ttc tat ttt ggg aag ctt gat ggc cag atc tcc tct			433
Lys Asp Arg Tyr Phe Tyr Phe Gly Lys Leu Asp Gly Gln Ile Ser Ser			
130	135	140	
gcc tac ccc agc caa gag ggg cag gtg ctg gtg ggc atc tat ggc cag			481
Ala Tyr Pro Ser Gln Glu Gly Gln Val Leu Val Gly Ile Tyr Gly Gln			
145	150	155	
tat caa ctc ctt ggc atc aag agc att ggc ttt gaa tgg aat tat cca			529
Tyr Gln Leu Leu Gly Ile Lys Ser Ile Gly Phe Glu Trp Asn Tyr Pro			
160	165	170	
cta gag gag ccg acc act gag cca cca gtt aat ctc aca tac tca gca			577
Leu Glu Glu Pro Thr Thr Glu Pro Pro Val Asn Leu Thr Tyr Ser Ala			
175	180	185	190
aac tca ccc gtg ggt cgc taggggtgggg tatggggcca tccgagctga ggcca			630
Asn Ser Pro Val Gly Arg			
195			
tcigtgtggt ggtggctgat ggtactggag taactgagtc gggacgctga atctgaatcc			690
accaataaat aaagcttctg c			711

【0 0 9 0】

<210> 28

<213> Homo Sapience

<220>

<221> CDS

<222> (242)...(565)

<400> 28

```

aaagaaacaa gccgggggac tgcgagccag ggactcgggc cgcggggcggaagaagtgg      60
ggcagcgctt ggccaggccg aaaggacttt gggggtgggg gctgggagtc cgtgtctcga      120
atgagggagg agaggtggag ttgccggggc tcaggcccgg cctcgagcat gggcggatga      180
gaggagtcgg gagccgaggc ctagggtcct tcgggtgagg ggagacggag ccagcgagga      240
g atg gag cag aag ctt gtg gag gag att ctt caa gca atc act atg      286
Met Glu Gln Lys Leu Val Glu Glu Ile Leu Gln Ala Ile Thr Met
      1              5              10             15
tca aca gac aca ggt gtt tcc ctt cct tca tat gag gaa gat cag gga      334
Ser Thr Asp Thr Gly Val Ser Leu Pro Ser Tyr Glu Glu Asp Gln Gly
      20              25              30
tca aaa ctc att cga aaa gct aaa gag gca cca ttc gta ccc gtt gga      382
Ser Lys Leu Ile Arg Lys Ala Lys Glu Ala Pro Phe Val Pro Val Gly
      35              40              45
ata gcg ggt ttt gca gca att gtt gca tat gga tta tat aaa ctg aag      430
Ile Ala Gly Phe Ala Ala Ile Val Ala Tyr Gly Leu Tyr Lys Leu Lys
      50              55              60
agc agg gga aat act aaa atg tcc att cat ctg atc cac atg cgt gtg      478
Ser Arg Gly Asn Thr Lys Met Ser Ile His Leu Ile His Met Arg Val
      65              70              75
gca gcc caa ggc ttt gtt gta gga gca atg act gtt ggt atg ggc tat      526
Ala Ala Gln Gly Phe Val Val Gly Ala Met Thr Val Gly Met Gly Tyr
      80              85              90              95
tcc atg tat cgg gaa ttc tgg gca aaa cct aag cct tagaagaa      570
Ser Met Tyr Arg Glu Phe Trp Ala Lys Pro Lys Pro

```

100

105

gagatgctgt cttggctctg ttggaggagc ttgctttagt tagatgtctt attattaaag 630
ttacctatta ttgttgaaa t 651

【 0 0 9 1 】

<210> 29

<211> 1310

<212> DNA

<213> Homo Sapience

<220>

<221> CDS

<222> (78)...(1130)

<400> 29

cgaacgccaa ggcggccacg tcctgctccc cctgggtgaag aagctgccct gggcttgtcg 60

tcctagggtc tccagac atg tct gag gtg aag agc cgg aag aag tgc ggg 110

Met Ser Glu Val Lys Ser Arg Lys Lys Ser Gly

1

5

10

ccc aag gga gcc cct gct gcg gag ccc ggg aag cgg agc gag ggc ggg 158

Pro Lys Gly Ala Pro Ala Ala Glu Pro Gly Lys Arg Ser Glu Gly Gly

15

20

25

aag acc ccc gtg gcc cgg agc agc gga ggc ggg ggc tgg gca gac ccc 206

Lys Thr Pro Val Ala Arg Ser Ser Gly Gly Gly Gly Trp Ala Asp Pro

30

35

40

cga acg tgc ctg agc ctg ctg tgc ctg ggg acg tgc ctg ggc ctg gcc 254

Arg Thr Cys Leu Ser Leu Leu Ser Leu Gly Thr Cys Leu Gly Leu Ala

45

50

55

tgg ttt gta ttt cag cag tca gaa aaa ttt gca aag gtg gaa aac caa 302

Trp Phe Val Phe Gln Gln Ser Glu Lys Phe Ala Lys Val Glu Asn Gln

Tyr Gln Leu Leu Lys Leu Glu Thr Asn Glu Phe Gln Gln Leu Gln Ser

80	85	90	
aaa atc agt tta att tca gaa aag tgg cag aaa tct gaa gct atc atg			398
Lys Ile Ser Leu Ile Ser Glu Lys Trp Gln Lys Ser Glu Ala Ile Met			
95	100	105	
gaa caa ttg aag tct ttt caa ata att gct cat cta aag cgt cta cag			446
Glu Gln Leu Lys Ser Phe Gln Ile Ile Ala His Leu Lys Arg Leu Gln			
110	115	120	
gaa gaa att aat gag gta aaa act tgg tcc aat agg ata act gaa aaa			494
Glu Glu Ile Asn Glu Val Lys Thr Trp Ser Asn Arg Ile Thr Glu Lys			
125	130	135	
cag gat ata ctg aac aac agt ctg acg acg ctt tct caa gac att aca			542
Gln Asp Ile Leu Asn Asn Ser Leu Thr Thr Leu Ser Gln Asp Ile Thr			
140	145	150	155
aaa gta gac caa agt aca act tcc atg gca aaa gat gtt ggt ctc aag			590
Lys Val Asp Gln Ser Thr Thr Ser Met Ala Lys Asp Val Gly Leu Lys			
160	165	170	
att aca agt gta aaa aca gat ata cga cgg att tca ggt tta gta act			638
Ile Thr Ser Val Lys Thr Asp Ile Arg Arg Ile Ser Gly Leu Val Thr			
175	180	185	
gat gta ata tca ttg aca gat tct gtg caa gaa cta gaa aat aaa ata			686
Asp Val Ile Ser Leu Thr Asp Ser Val Gln Glu Leu Glu Asn Lys Ile			
190	195	200	
gag aaa gta gaa aaa aat aca gta aaa aat ata ggt gat ctt ctt tca			734
Glu Lys Val Glu Lys Asn Thr Val Lys Asn Ile Gly Asp Leu Leu Ser			
205	210	215	
agc agt att gat cga aca gca acg ctc cga aag aca gca tct gaa aat			782
Ser Ser Ile Asp Arg Thr Ala Thr Leu Arg Lys Thr Ala Ser Glu Asn			
220	225	230	235
tca caa aga att aac tct gtt aag aag acg cta acc gaa cta aag agt			830

Ser Gln Arg Ile Asn Ser Val Lys Lys Thr Leu Thr Glu Leu Lys Ser	
240	245
gac ttc gac aaa cat aca gat aga ttt cta agc tta gaa ggt gac aga	878
Asp Phe Asp Lys His Thr Asp Arg Phe Leu Ser Leu Glu Gly Asp Arg	
255	260
gcc aaa gtt ctg aag aca gtg act ttt gca aat gat cta aaa cca aag	926
Ala Lys Val Leu Lys Thr Val Thr Phe Ala Asn Asp Leu Lys Pro Lys	
270	275
gtg tat aat cta aag aag gac ttt tcc cgt tta gaa cca tta gta aat	974
Val Tyr Asn Leu Lys Lys Asp Phe Ser Arg Leu Glu Pro Leu Val Asn	
285	290
gat tta aca cta cgc att ggg aga ttg gtt acc gac tta cta caa aga	1022
Asp Leu Thr Leu Arg Ile Gly Arg Leu Val Thr Asp Leu Leu Gln Arg	
300	305
gag aaa gaa att gct ttc tta agt gaa aaa ata tct aat tta aca ata	1070
Glu Lys Glu Ile Ala Phe Leu Ser Glu Lys Ile Ser Asn Leu Thr Ile	
320	325
gtc caa gct gag att aag gat att aaa gat gaa ata gca cac att tca	1118
Val Gln Ala Glu Ile Lys Asp Ile Lys Asp Glu Ile Ala His Ile Ser	
335	340
gat atg aat tagtttgaca ttattgagat tagactaagg taattttttt aat	1170
Asp Met Asn	
350	
gggacctctc atgagaagac tggtaaataca aaaataatga ttttttgag caaaagtcac	1230
tttatattta atcctatttt gtacagtaaa aataaaactt taaaacaggt tgattttcca	1290
aaataaatat gctaaaacct	1310

<212> DNA

<213> Homo Sapience

<220>

<221> CDS

<222> (233)...(556)

<400> 30

tggtgtatg ctattggagg gtggaaatca catctcctgt ttatccgtgt gcttgtagg	60
tgtcagccgc ccccccccc ccatatgcag atttactcgg catggtagtg gccagcttct	120
aacacagctg gtatttcaag tctcctggga cctcactcag gaatgatacc ccctcagtag	180
aagcagcagg tgatcttaac tcctttcaaa gagcaggcct gtctgggaag cc atg	235

Met

1

tcc tca gca ggc aca gca acc cct ctg gaa atg gat cac aaa ctc act	283
Ser Ser Ala Gly Thr Ala Thr Pro Leu Glu Met Asp His Lys Leu Thr	

5

10

15

tct cag cca ggc agg cca agc ttc tat tgt aac agt agg cac agt ata	331
Ser Gln Pro Gly Arg Pro Ser Phe Tyr Cys Asn Ser Arg His Ser Ile	

20

25

30

gtc gga tca tca cat cag ctg ggt ttt tgg ttt agt cat cta gag tcg	379
Val Gly Ser Ser His Gln Leu Gly Phe Trp Phe Ser His Leu Glu Ser	

35

40

45

tct gga cta aag gtc ttt cag gtc tcc ttg ccc tgt gag tgc gtg aac	427
Ser Gly Leu Lys Val Phe Gln Val Ser Leu Pro Cys Glu Cys Val Asn	

50

55

60

65

ctc ccc acc cga att gcc tca gtt gtc ctg agc ctc atg tct ctc ctg	475
Leu Pro Thr Arg Ile Ala Ser Val Val Leu Ser Leu Met Ser Leu Leu	

70

75

80

gtg gtg ggc cag gcc cct gca tgg gaa ggg agc ctg ctg cgg ggc agg	523
Val Val Gly Gln Ala Pro Ala Trp Glu Gly Ser Leu Leu Arg Gly Arg	

85	90	95	
cca gct ggg ggt gct cac cta tgc gca gca tgaagttatt gaaggac			570
Pro Ala Gly Gly Ala His Leu Cys Ala Ala			
100	105		
tggttggttga tggttggtgag cgtatccttc atggccagcg cgaagtcggc caggtcagcc			630
aggtgctgcc agcgctctct ctcggacttg tcttcctgtg ccaggggacc gtggagaaag			690
tgtcaggggc cgtcactgc agcagcctgc tctgctgcct tccctggcag tgttctgggg			750
gtggattccc tacacctaga tgttcaaggc cttacttttc ctcccacaaa ggagtcgcag			810
ccacgctagc tctgacttgc cactgtgaca aagttcacgt agcaggtcta ggcaaagact			870
gggcaattga gcagaggaga cggacctgtg agtctgacca cgaggcggac cccttcacct			930
tggctggggc tggctcctggt ccttaggtti tgcaggilg tccctgtttg gatccctcaa			990
ctaggigata agcactggag ggggatgacc cgccttggac gtgtttcttt aacctcatcc			1050
atataatagg gccgtgggat ggttgttagag gtaaagcagg atgatggtgt tttaagacca			1110
gagcttggga ccagggctcc tacacctaat ttctctcct ggtagctgaa caaaggtcta			1170
aattagctta acaaaagaac aggtgcccgt cagccagagt tctgaaggcc atgctttcag			1230
tttcccttgt tgacaattgc tctccagttc ctatgaaagc acagagcctt agggggcctg			1290
gccacagaac acaaccatct taggcctgag ctgtgaacag caggggggttg tgtgtctgtt			1350
ctgtttctct gcttgccgaa ctttctcaat aaacctatt tcttatttat			1400

【図面の簡単な説明】

【図1】 クローンHP01470がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図2】 クローンHP02419がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図3】 クローンHP02631がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図4】 クローンHP02695がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

図1、図2、図3、図4は、それぞれ、クローンHP01470、クローンHP02419、クローンHP02631、クローンHP02695がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図 6】 クローン HP 10530 がコードする蛋白質の疎水性／親水性プロフィールを示す図である。

【図 7】 クローン HP 10541 がコードする蛋白質の疎水性／親水性プロフィールを示す図である。

【図 8】 クローン HP 10550 がコードする蛋白質の疎水性／親水性プロフィールを示す図である。

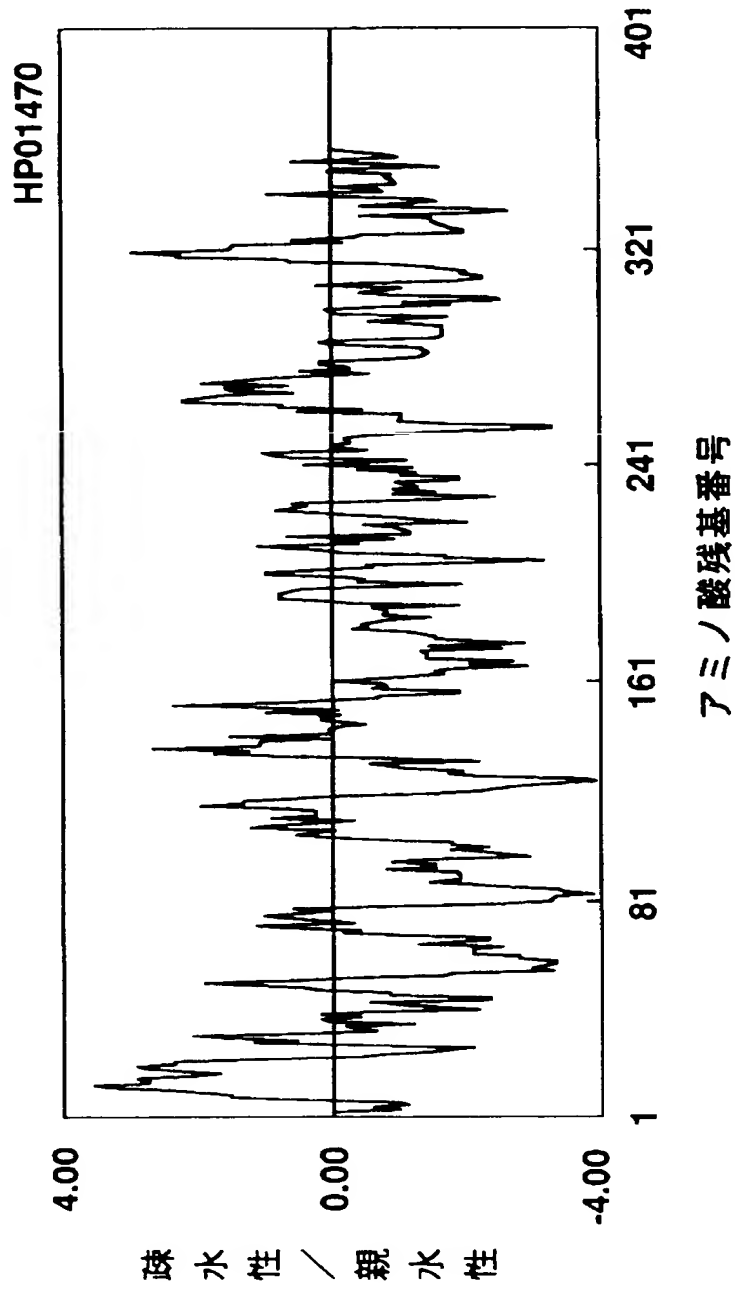
【図 9】 クローン HP 10590 がコードする蛋白質の疎水性／親水性プロフィールを示す図である。

【図 10】 クローン HP 10591 がコードする蛋白質の疎水性／親水性プロフィールを示す図である。

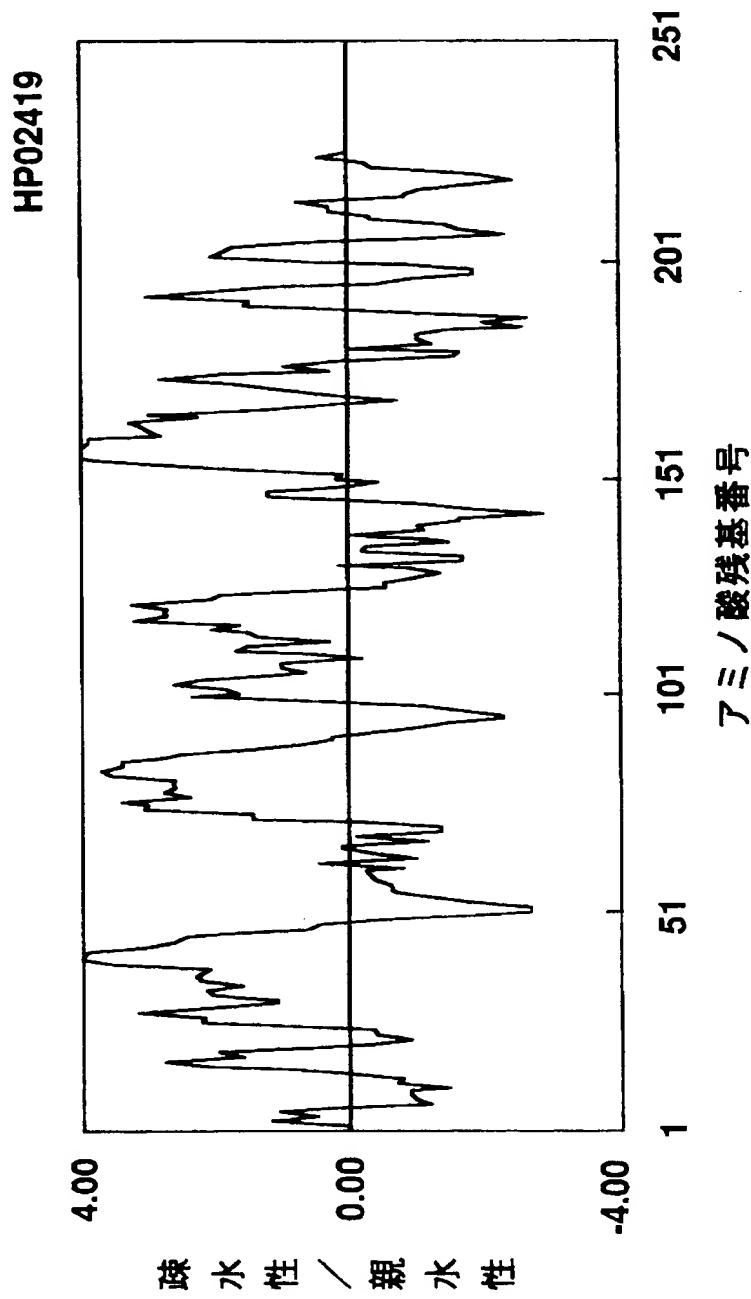
【書類名】

図面

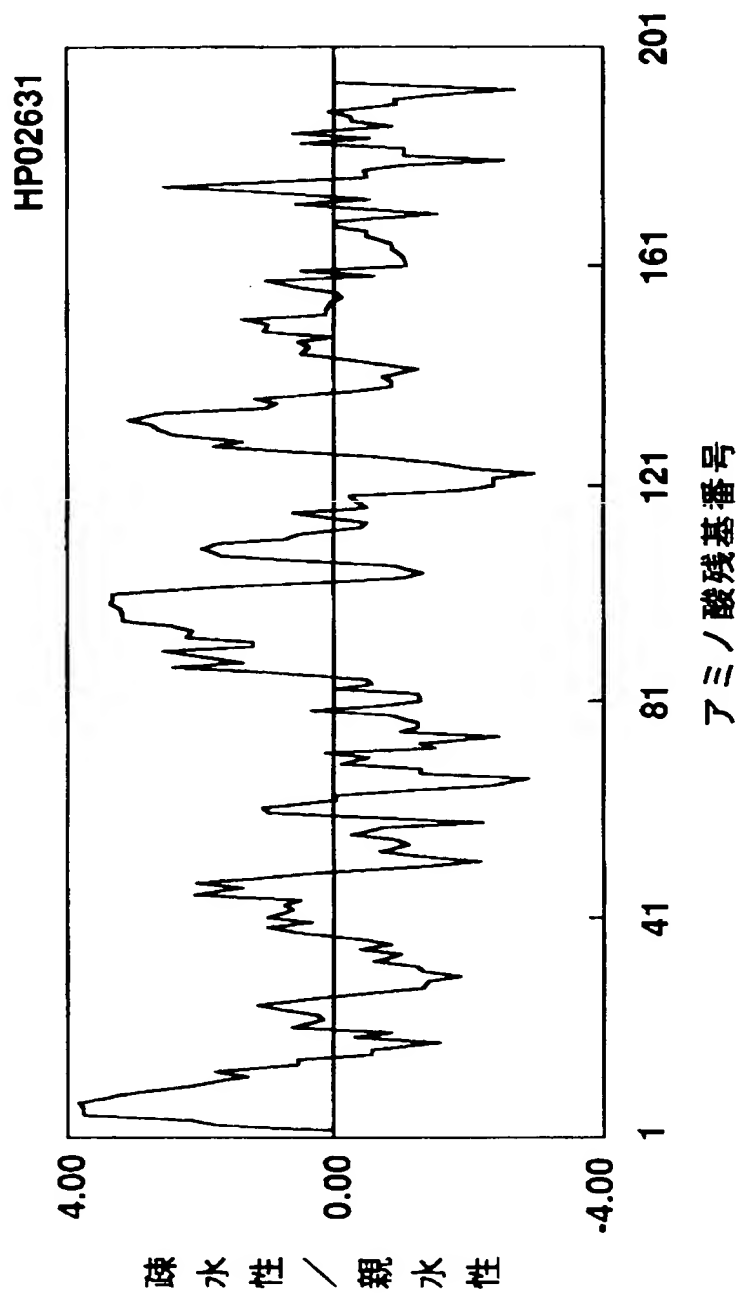
【図1】



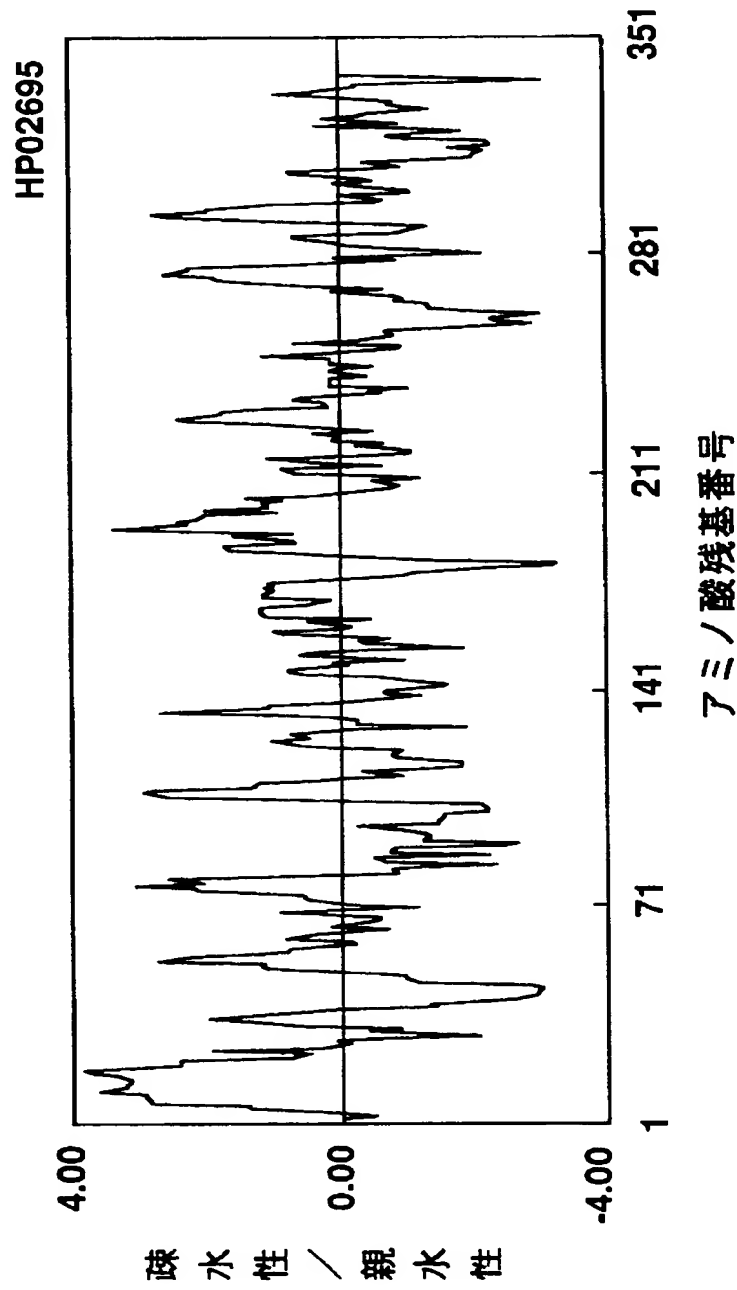
【図 2】



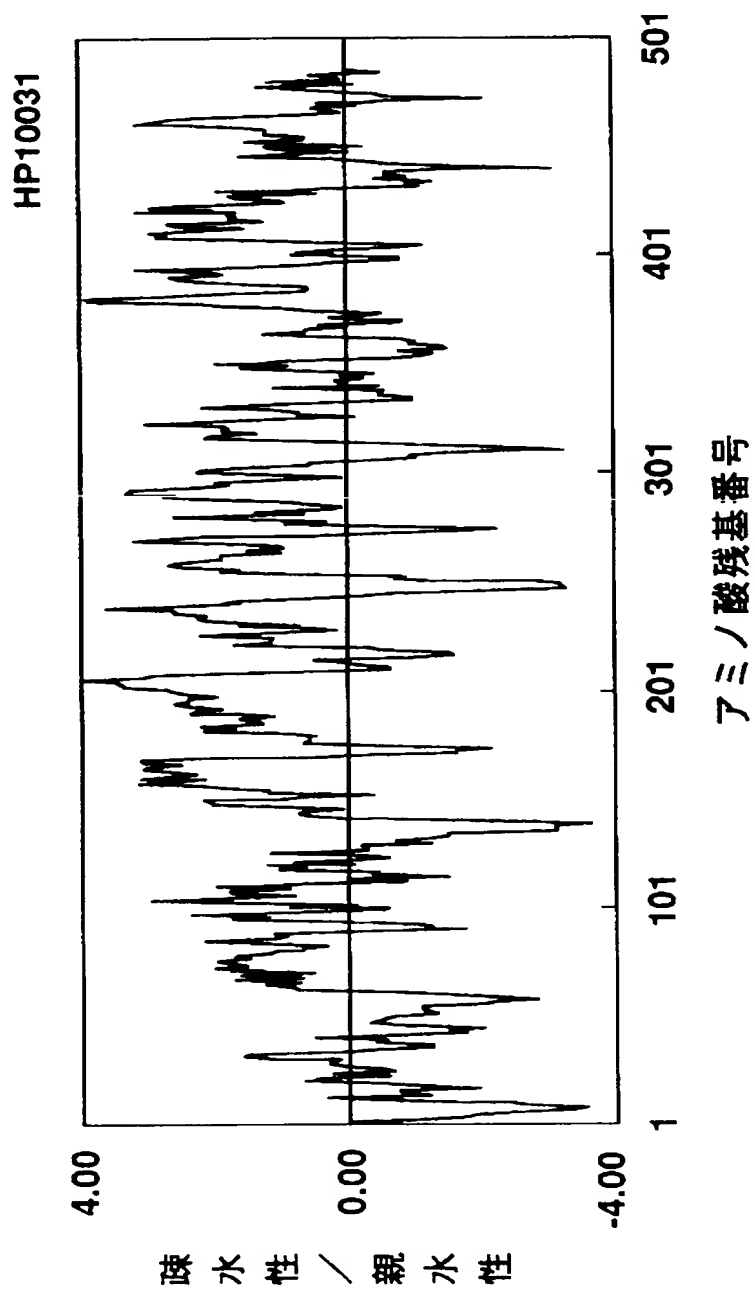
【図 3】



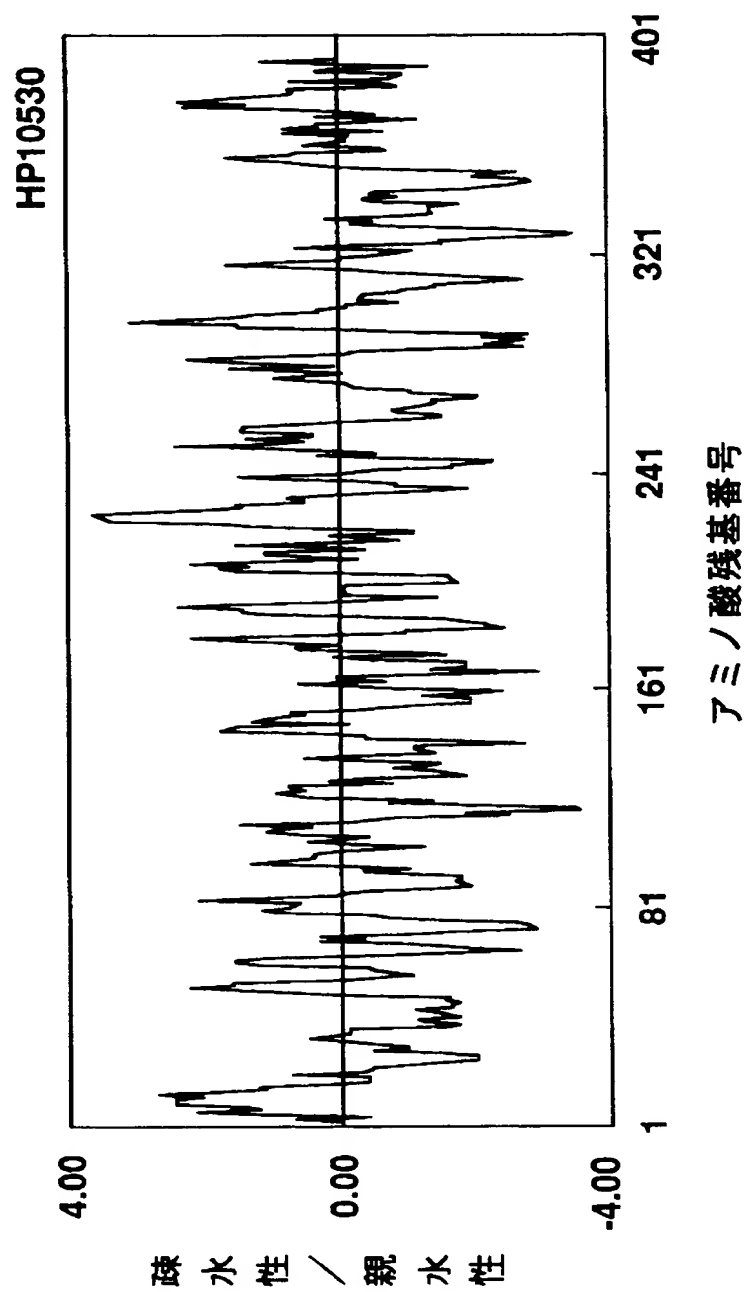
【図 4】



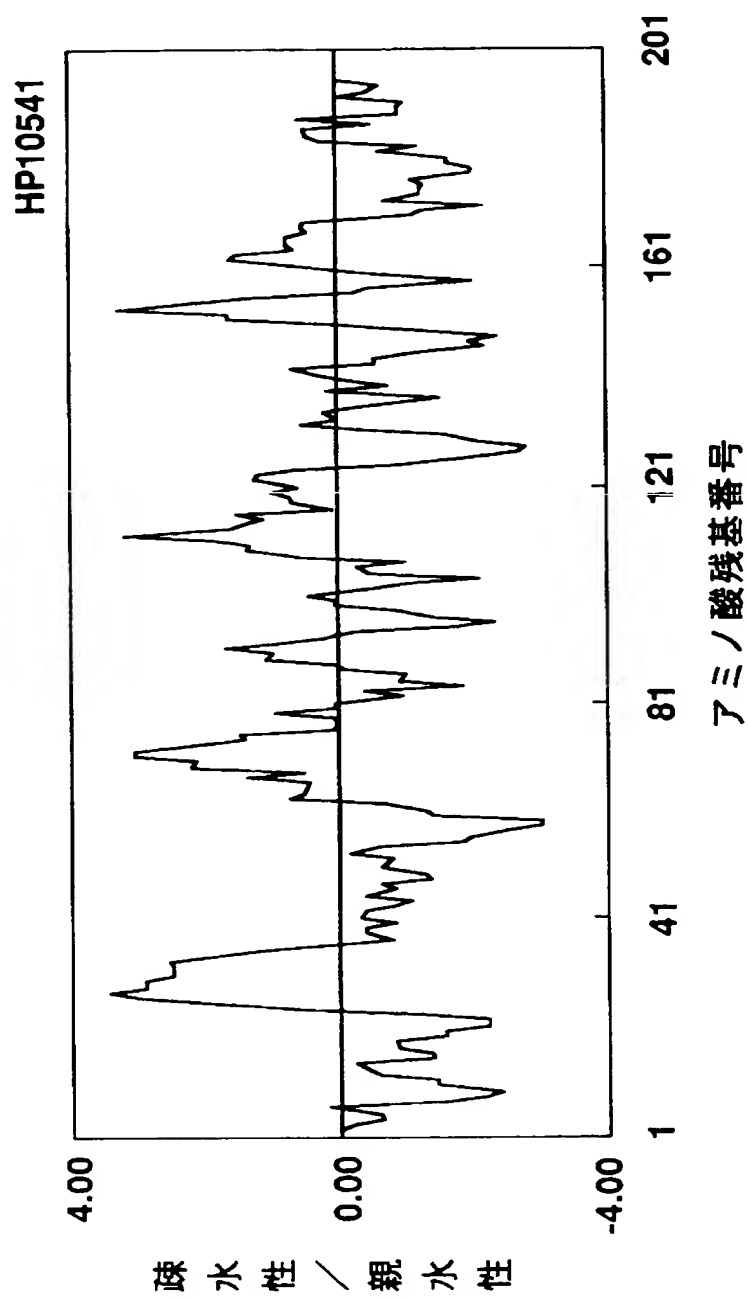
【図5】



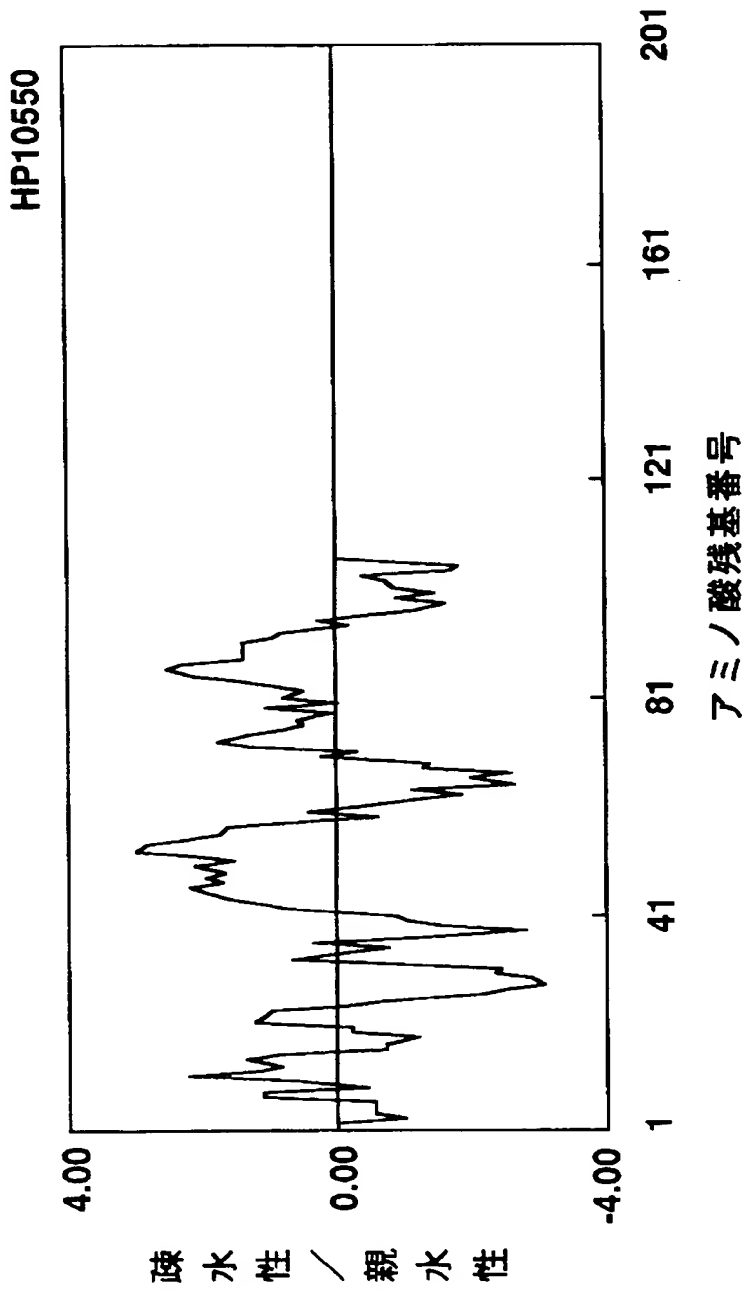
【図 6】



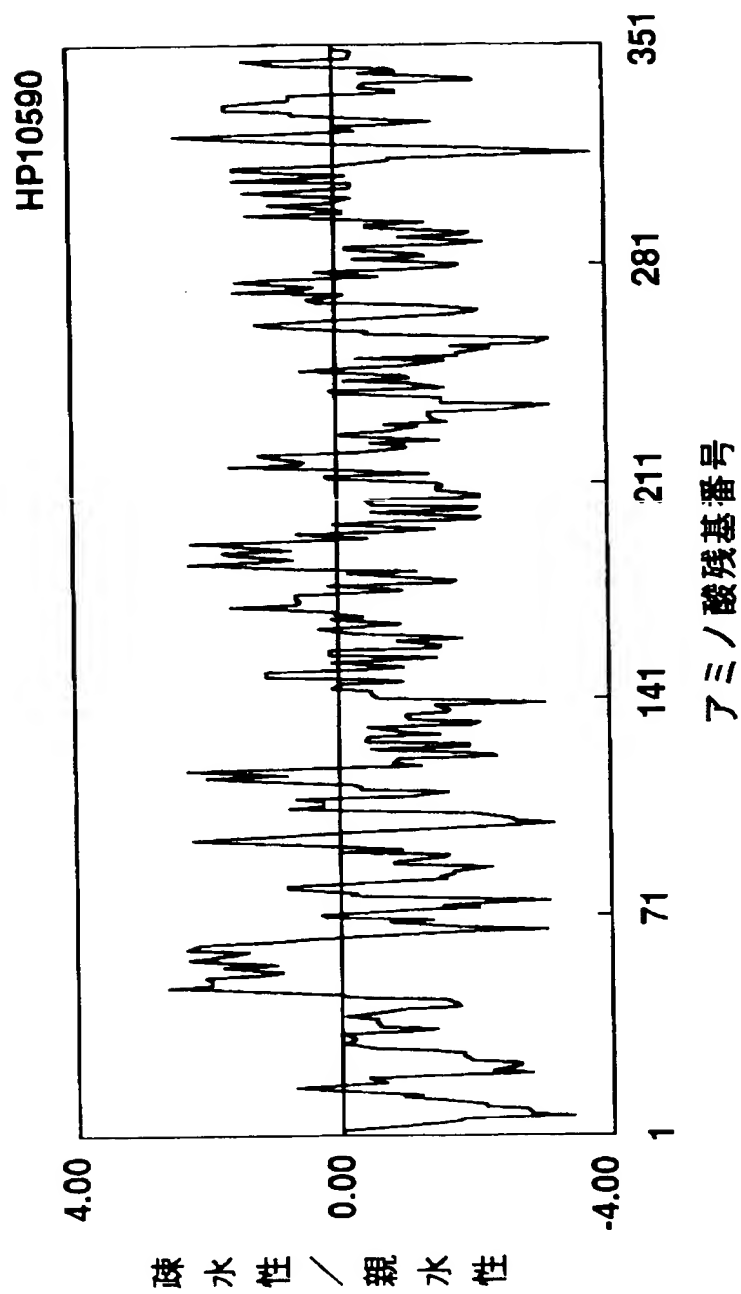
【図7】



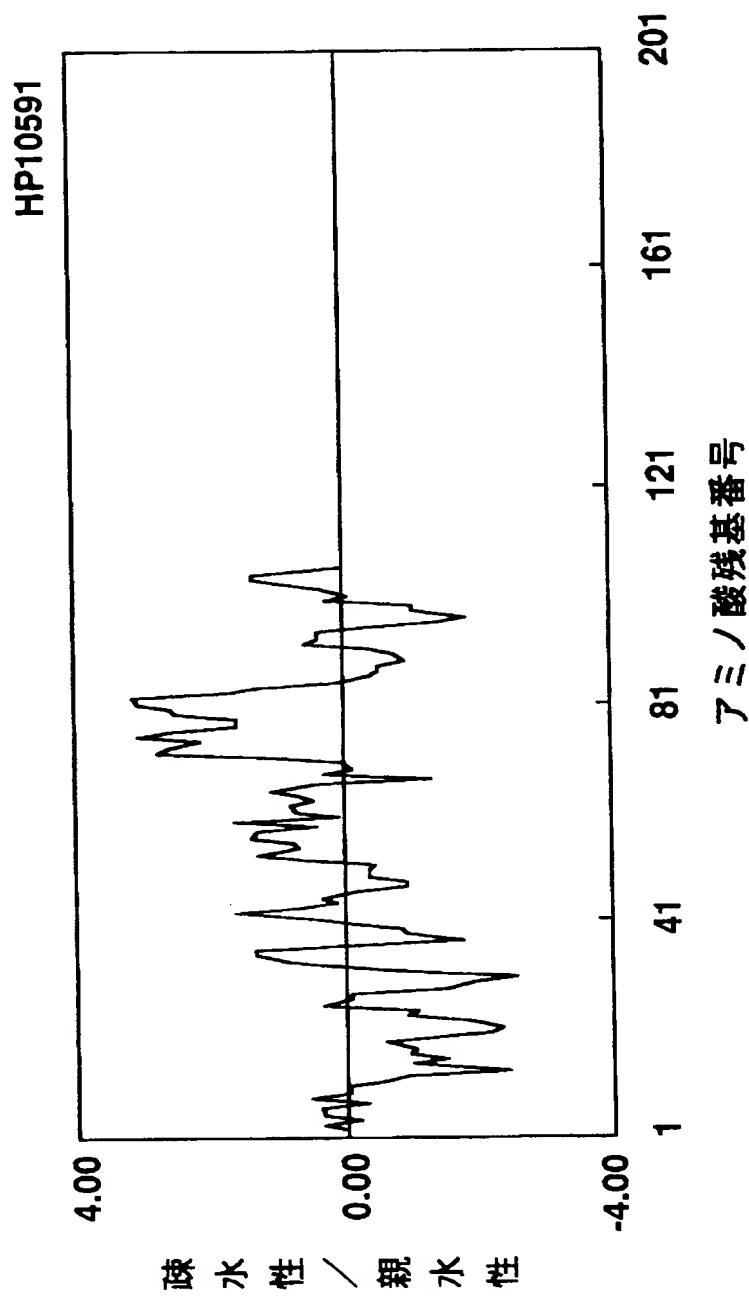
【図 8】



【図9】



【図 10】



【書類名】 要約書

【要約】

【課題】 疎水性ドメインを有するヒト蛋白質、それをコードしている cDNA、この cDNA の発現ベクター、およびこの cDNA を発現させた真核細胞を提供する。

【解決手段】 配列番号 1 から配列番号 10 で表されるアミノ酸配列のいずれかを含む蛋白質、この蛋白質をコードする DNA、例えば配列番号 11 から配列番号 20 で表される塩基配列を含む cDNA、この cDNA の発現ベクター、およびこの cDNA を発現させた真核細胞。疎水性ドメインを有するヒト蛋白質をコードしている cDNA の組換え体を発現させることにより、この蛋白質並びにこの蛋白質を発現する真核細胞を提供することができる。

【選択図】 なし

【書類名】
【訂正書類】

職権訂正データ
特許願

<認定情報・付加情報>

【特許出願人】

申請人

【識別番号】

000173762

【住所又は居所】

神奈川県相模原市西大沼4丁目4番1号

【氏名又は名称】

財団法人相模中央化学研究所

【特許出願人】

【識別番号】

596134998

【住所又は居所】

東京都目黒区中町2丁目20番3号

【氏名又は名称】

株式会社プロテジーン

出 願 人 履 歴 情 報

識別番号 [000173762]

1. 変更年月日 1995年 4月14日
[変更理由] 住所変更
住 所 神奈川県相模原市西大沼4丁目4番1号
氏 名 財団法人相模中央化学研究所

出 願 人 履 歴 情 報

識別番号

[596134998]

1. 変更年月日

1996年 9月13日

[変更理由]

新規登録

住 所

東京都目黒区中町2丁目20番3号

氏 名

株式会社プロテジーン